



Au-delà de la mesure de SUV en imagerie TEP : propriétés et potentiel des paramètres de texture pour caractériser les tumeurs

Fanny Orlhac

► To cite this version:

Fanny Orlhac. Au-delà de la mesure de SUV en imagerie TEP : propriétés et potentiel des paramètres de texture pour caractériser les tumeurs. Physique Médicale [physics.med-ph]. Université Paris Sud - Paris XI, 2015. Français. NNT : 2015PA112198 . tel-01390197

HAL Id: tel-01390197

<https://theses.hal.science/tel-01390197>

Submitted on 1 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Comprendre le monde,
construire l'avenir®



UNIVERSITÉ PARIS-SUD

**ÉCOLE DOCTORALE : Sciences et Technologie de l'Information,
des Télécommunications et des Systèmes**

Laboratoire Imagerie Moléculaire In Vivo (IMIV)
UMR 1023 Inserm/CEA-SHFJ/Université Paris Sud - ERL 9218 CNRS

DISCIPLINE : Physique

THÈSE DE DOCTORAT

soutenue le 22/09/2015

par

Fanny ORLHAC

AU-DELÀ DE LA MESURE DE SUV EN IMAGERIE TEP :
*propriétés et potentiel des paramètres de texture
pour caractériser les tumeurs*

Directeur de thèse : Irène BUVAT Directeur de Recherche (CNRS, Université Paris-Sud)

Composition du jury :

Président : Eric DEUTSCH PU-PH (Université Paris-Sud)

Rapporteurs : Jacques DARCOURT PU-PH (Université de Nice)
Nicolas PASSAT Professeur (Université de Reims Champagne-Ardenne)

Examineurs : Carole LARTIZIEN Chargée de Recherche (CNRS, Université Lyon 1)
Samuel VALABLE Chargé de Recherche (CNRS, Université de Caen)

Membre invité : Michaël SOUSSAN MCU-PH (Université Paris 13)

Remerciements

J'adresse tout d'abord mes plus vifs remerciements à Irène Buvat pour sa confiance et son encadrement efficace durant ces trois années. Merci d'avoir toujours été disponible, de ne jamais avoir laissé la moindre de mes questions sans réponse et de m'avoir guidée par de précieux conseils tout au long de ce travail. Tous les thésards n'ont pas cette chance.

Je remercie Jacques Darcourt et Nicolas Passat d'avoir accepté d'être rapporteurs de cette thèse. Leur relecture attentive et leurs remarques ont permis d'améliorer ce manuscrit. Je remercie également Eric Deutsch, Carole Lartizien et Samuel Valable pour leur participation à mon jury de thèse et l'intérêt qu'ils ont porté à ce travail.

Mes remerciements vont également à Michaël Soussan, mon « co-thésard ». Notre collaboration a été fructueuse et très enrichissante. Ce travail d'équipe a toujours été très agréable grâce à sa bonne humeur et sa persévérance.

Merci à Christophe Nioche qui a accepté de partager ses données, pour ses conseils sur les contours ô combien complexes des gliomes, et pour avoir été un très bon compagnon de voyage à Baltimore. Je sais que le logiciel est entre de bonnes mains.

Merci à Benoit Thézé d'avoir réalisé, avec une très grande minutie, la partie bio de l'étude multi-échelle.

J'adresse également mes remerciements aux services de médecine nucléaire de l'hôpital Avicenne de Bobigny, de l'hôpital d'Instruction des Armées du Val de Grâce et de l'Institut Jules Bordet de Bruxelles, en particulier à Camilo Garcia, pour leurs données cliniques, indispensables afin de tester les développements méthodologiques.

Merci à Charlotte et Morgane pour leur accompagnement pendant cette thèse. Nos discussions, leur présence et leur grande humanité ont été un soutien indispensable, merci pour votre amitié.

Je remercie également mes collègues docteurs et futurs docteurs : Nicolas, Didier, Vincent, Pauline, Yann, Cécile, Sara, Amandine, Emilie et Hayet. Bonne chance à tous pour la suite. Merci à Jacques-Antoine pour son encadrement pendant mon stage de M2. Je n'oublie pas non plus les personnels scientifiques et administratifs des laboratoires IMNC et IMIV, et, en particulier, les membres des équipes QIM et de Physique Biomédicale.

Merci à mes collègues du département de Mesures Physiques de l'IUT d'Orsay pour leurs conseils lors de mes premiers pas d'enseignante.

Parce que certaines rencontres ne s'oublient pas, merci à Mireille Baudino qui m'a transmis le goût de la physique (en me détournant un peu des maths !) il y a très longtemps... Merci à Hélène Per-

rin d'avoir été la première à m'ouvrir les portes d'un laboratoire de recherche. Je garde un excellent souvenir de ce stage à Villetaneuse.

Enfin, je termine en remerciant ma famille. Merci à mes parents d'avoir toujours appuyé mes choix d'études, d'avoir été un soutien moral, affectif, matériel, financier et culinaire durant toutes ces années. J'ai toujours été portée par leurs encouragements et leur confiance. Je leur dédie ce travail. Merci également à mon « fan club » de la première heure, vous avez été d'un grand soutien et même si vous n'arrivez pas à la fin du manuscrit, je ne vous en voudrai pas !

Table des matières

Résumé	1
Abstract	3
Introduction	5
1 La tomographie par émission de positons en oncologie	7
1.1 Imagerie TEP en clinique	7
1.1.1 Principe de l'imagerie TEP	7
1.1.2 Imagerie TEP au 18F-FDG	12
1.1.3 Autres traceurs marqués au 18F en oncologie	13
1.2 Analyse des images TEP	15
1.2.1 Estimation de l'activité métabolique des tumeurs	15
1.2.2 Estimation du volume métabolique	19
1.2.3 Caractérisation tumorale	21
1.3 Place de la TEP au 18F-FDG dans la prise en charge en oncologie	27
1.3.1 Diagnostic et bilan d'extension	27
1.3.2 Planification de traitement en radiothérapie	27
1.3.3 Évaluation de la réponse aux traitements	28
1.3.4 Surveillance des récives	31
2 Caractérisation de l'hétérogénéité tumorale	33
2.1 Motivations	33
2.1.1 Oncogenèse	33
2.1.2 Hétérogénéité tumorale	34
2.1.3 Nécessité d'identifier des biomarqueurs	36
2.2 Principe de l'analyse de texture	36
2.2.1 Protocole de calcul	36
2.2.2 Définition des matrices de texture	37
2.3 État de l'art en imagerie TEP	43
2.3.1 Introduction	43
2.3.2 Comparaison des choix méthodologiques pour le calcul des index	48
2.3.3 Robustesse des index	49
2.3.4 Analyse des études cliniques	52
3 Étude méthodologique des paramètres de texture	57
3.1 Matériels & méthodes	57
3.1.1 Patients	57
3.1.2 Acquisition des données TEP/TDM	58
3.1.3 Segmentation du volume métabolique	59

3.1.4	Analyse des images	59
3.2	Étude de la corrélation entre index	60
3.2.1	Matériels & méthodes	60
3.2.2	Résultats	60
3.2.3	Discussion	61
3.3	Impact de la méthode de segmentation	63
3.3.1	Matériels & méthodes	63
3.3.2	Résultats	64
3.3.3	Discussion	64
3.4	Impact de l'étape de discrétisation	68
3.4.1	Taux de discrétisation	68
3.4.2	Méthode de discrétisation relative	70
3.4.3	Méthode de discrétisation absolue	75
3.5	Interprétation des paramètres de texture en TEP	85
3.5.1	Matériels & méthodes	85
3.5.2	Résultats	86
3.5.3	Discussion	90
4	Évaluation clinique des index de texture	97
4.1	Distinction entre les tissus sains et tumoraux	97
4.1.1	Matériels & méthodes	97
4.1.2	Résultats	98
4.1.3	Discussion	101
4.2	Corrélation avec l'analyse visuelle	107
4.2.1	Matériels & méthodes	107
4.2.2	Résultats	107
4.2.3	Discussion	112
4.3	Corrélation avec les caractéristiques tumorales	115
4.3.1	Les lésions mammaires	115
4.3.2	Les lésions pulmonaires non à petites cellules	120
4.3.3	Les gliomes	127
5	Étude multi-échelle de l'hétérogénéité tumorale	137
5.1	Matériels & méthodes	137
5.1.1	Modèles animaux	137
5.1.2	Protocole de l'imagerie TEP/TDM	137
5.1.3	Autoradiographie	138
5.1.4	Histochimie	138
5.1.5	Préparation des volumes 3D	138
5.1.6	Analyse des images	141
5.2	Impact de la taille des voxels et de la résolution spatiale	141
5.2.1	Matériels & méthodes	141
5.2.2	Résultats	141
5.2.3	Discussion	142
5.3	Comparaison de l'analyse de texture à différentes échelles	146
5.3.1	Matériels & méthodes	146
5.3.2	Résultats	146
5.3.3	Discussion	147
5.4	Impact de la densité cellulaire sur l'analyse de texture	150
5.4.1	Matériels & méthodes	150

5.4.2	Résultats	151
5.4.3	Discussion	152
Discussion générale		157
Conclusion et perspectives		163
Bibliographie		167
A	Liste des abréviations	179
B	Liste des publications	183

Résumé

Caractériser précisément l'hétérogénéité tumorale constitue un enjeu majeur en cancérologie. En effet, la coexistence de différentes populations de cellules affecte la croissance tumorale et donc la stratégie thérapeutique à adopter. Le calcul de biomarqueurs de cette hétérogénéité directement à partir des données d'imagerie présente de nombreux avantages : il est non-invasif, répétable plusieurs fois au cours du traitement, ne nécessite pas d'examen supplémentaire et permet de caractériser la tumeur toute entière et ses éventuelles métastases, à la différence des biopsies. En imagerie TEP, au-delà de l'interprétation classique par des mesures de fixation normalisées (SUV), de volume métabolique ou de glycolyse de la lésion, l'utilisation de paramètres de texture a été proposée pour caractériser l'hétérogénéité tumorale. Depuis 2009, des études ont montré que ces index, extraits des images TEP avant traitement, permettaient de prédire la réponse à la thérapie et/ou la survie du patient pour de nombreux types de cancer.

Mon projet de recherche visait à développer et évaluer des méthodes pour une caractérisation plus complète de l'activité métabolique des tumeurs.

L'analyse de texture des images TEP nécessite un protocole de calcul des index plus complexe que celui des paramètres conventionnels utilisés en clinique. Afin de déterminer l'influence des étapes préliminaires au calcul de ces index, une étude méthodologique a tout d'abord été réalisée sur quatre groupes de patients. Cette analyse a montré que certains paramètres de texture étaient redondants et qu'il existait une forte corrélation entre certains d'entre eux et le volume métabolique. Elle a également mis en évidence l'impact de la formule et du taux de discrétisation sur les valeurs des paramètres de texture et permis de clarifier l'interprétation des indices.

Après avoir établi un protocole de calcul strict, une seconde partie de ce travail a consisté à évaluer la capacité de ces index pour la caractérisation des tumeurs. L'analyse de texture a ainsi permis de différencier les tissus sains des tissus tumoraux dans le cas de cancers pulmonaires ou mammaires, et d'identifier les tumeurs primitives des tumeurs récurrentes chez des patients atteints de gliomes. Nous avons également montré que les paramètres de texture permettaient de distinguer les types histologiques comme les tumeurs mammaires triple-négatives des non triple-négatives, les lésions pulmonaires épidermoïdes des adénocarcinomes ou encore les gliomes de haut grade des lésions de bas grade.

Afin de comprendre le lien entre l'hétérogénéité tumorale quantifiée sur les images TEP et l'hétérogénéité biologique des lésions, nous avons comparé l'analyse de texture réalisée à différentes échelles sur un modèle animal. Cette étude a révélé que la texture mesurée *in vivo* sur les images TEP reflétait la texture mesurée *ex vivo* sur les images autoradiographiques. L'analyse des coupes histologiques a mis en évidence que les paramètres extraits des images TEP ne reflétaient pas point par point l'hétérogénéité de distribution des cellules tumorales mais plutôt l'organisation générale de ces cellules dans la lésion.

A l'issue de ce travail, Entropy apparaît comme l'un des index de texture les plus prometteurs tant pour sa robustesse que ses capacités de caractérisation de l'hétérogénéité tumorale. Dans certains cas, Entropy conduit même à de meilleures performances que les index conventionnels. Des analyses complémentaires doivent être menées pour permettre son utilisation en clinique et mieux élucider sa signification biologique.

Mots clés : TEP, quantification, hétérogénéité tumorale, analyse de texture.

Abstract

The precise characterization of the biological heterogeneity of a tumor is a major issue in oncology. Indeed, the coexistence of different cell populations influences the tumor growth and the therapeutic strategy to be applied. The calculation of biomarkers reflecting this heterogeneity directly from imaging data offers a number of advantages: it is non-invasive, can be repeated during the therapy, does not require supplementary examinations and the whole tumor and possible metastases can be investigated from the images. In PET imaging, beyond the conventional interpretation based on the measurement of Standardized Uptake Value (SUV), metabolic volume or total lesion glycolysis, the use of texture parameters has been proposed. Since 2009, several studies have suggested that these indices, computed from PET images before treatment, could predict therapy response and/or patient survival for many different types of cancer.

My research project was to develop and assess methods to characterize the metabolic activity distribution in tumors.

Texture analysis based on PET images requires a protocol to compute index that is somehow more sophisticated than when simply measuring the conventional index used in clinical practice. To determine the role of the different steps that are involved in the computation of texture index, a methodological study was conducted using four patient groups. This study demonstrated that some texture parameters were redundant and that there existed a strong correlation between some of them and the metabolic volume. We have also shown that the formula and the rate of discretization impact the texture analysis and clarified the interpretation of these metrics.

After the protocol of texture index computation has been established, the second part of this work was to assess the interest of these indices for the tumor characterization. We showed that some texture indices were different in tumor and in healthy tissue in lung and breast cancers and were also different in recurrent tumors and in newly diagnosed tumors high grade gliomas. We demonstrated that using texture parameters, we could identify histological types such as the triple-negative breast tumors, the squamous cell carcinoma from adenocarcinoma in lung tumors, as well as the grade of gliomas.

To understand the links between the tumor heterogeneity as measured from PET images and the biological heterogeneity of lesions, we compared the texture analysis based on different scales in a mouse model. This study revealed that the texture measured *in vivo* based on PET images reflects the texture measured *ex vivo* from autoradiographic images. The analysis of histological slides indicates that the texture parameters measured from PET images did not reflect precisely the distribution of tumor cells, but rather characterize the general organization of cells in the lesion.

Based on our results, Entropy appears as one of the most promising texture indices thanks to its robustness and its ability to characterize the heterogeneity of tumor uptake. In some studies we conducted, Entropy even led to better performances than the conventional indices. Further research is still needed to make it usable in clinical practice and to better elucidate its biological meaning.

Keywords: PET, quantification, tumor heterogeneity, texture analysis.

Introduction

Proposer pour chaque patient un traitement spécifique afin d'améliorer la réponse à la thérapie, accroître sa survie tout en limitant les effets secondaires, tels sont les objectifs de la médecine personnalisée. Cette évolution de la cancérologie repose sur les progrès récents de la compréhension des processus tumoraux et le développement de thérapies ciblées. La mise en place de ces traitements « sur-mesures » nécessite en contrepartie une caractérisation précise de chaque tumeur. En clinique, quand cela est possible, une biopsie est réalisée avant le début du traitement afin d'identifier les spécificités de la lésion : présence ou non de certaines protéines ou récepteurs, voire identification du profil génomique ou protéomique de la tumeur. Cependant, ce profil ne caractérise que la région échantillonnée par la biopsie, limitée spatialement et temporellement, et ne rend pas compte de l'intégralité de la lésion, possiblement hétérogène.

En parallèle, l'imagerie médicale est devenue un outil indispensable en cancérologie et le développement des technologies n'a fait qu'accroître sa place dans la prise en charge du patient. Elle offre de nombreux avantages : fournir des informations sur la lésion entière, de façon non-invasive et répétable plusieurs fois au cours du traitement. Toutefois, dans la plupart des cas, l'analyse des images se limite à un repérage visuel de la position et de l'extension des tumeurs, voire une caractérisation assez sommaire de son activité métabolique.

En 2012, [Lambin et al., 2012, Kumar et al., 2012] ont présenté un concept nommé « radiomique » (ou « *radiomics* » en anglais) qui part du constat que les images médicales sont sous-exploitées. Les auteurs font l'hypothèse que les images renferment beaucoup plus d'informations que celles qui sont actuellement extraites et interprétées. Ils proposent de calculer, à partir des examens d'imagerie, plus de 200 index décrivant la distribution d'intensité du signal, la texture et la forme de la lésion ainsi que son interaction avec les tissus environnants, et de sélectionner les index les plus robustes. Ces index sont ensuite analysés conjointement avec la réponse aux traitements ou les caractéristiques génétiques des lésions et évalués pour leur valeur ajoutée par rapport à la stadification des cancers et aux données utilisées en clinique. L'hypothèse centrale de ce concept repose sur le fait que les caractéristiques génomiques et protéomiques des lésions peuvent s'exprimer en termes de caractéristiques macroscopiques mesurables sur les images de tomodensitométrie (TDM), d'imagerie par résonance magnétique (IRM) ou de tomographie par émission de positons (TEP).

En particulier, à l'opposé des modalités d'imagerie anatomique, l'imagerie TEP permet de visualiser, en 3 dimensions, dans le patient, la répartition d'un traceur ciblant un processus biologique d'intérêt. Largement utilisée en oncologie, cette imagerie fonctionnelle permet, grâce au traceur 2-desoxy-2-18F-fluoro-D-glucose (18F-FDG), de visualiser le métabolisme glucidique exacerbé des cellules de la plupart des tumeurs. Cet examen peut être utilisé comme outil de diagnostic, afin de guider la radiothérapie, pour évaluer la réponse aux traitements ou surveiller d'éventuelles récives. En effet, par rapport aux modalités d'imagerie anatomique, les examens TEP permettent de quantifier des modifications fonctionnelles associées à l'évolution tumorale et qui apparaissent préalablement aux modifications anatomiques.

En première intention, les images TEP sont analysées visuellement par le médecin nucléaire pour repérer le nombre et la localisation des lésions cancéreuses. Dans un second temps, une analyse quantitative peut être réalisée pour caractériser l'activité tumorale soit à partir de la mesure de valeur de fixation normalisée ou SUV (*Standardized Uptake Value*), soit à partir de la mesure du volume métaboliquement actif. Au-delà de ces mesures « classiques », de nouveaux paramètres issus de l'analyse de texture du signal ont été proposés pour quantifier plus précisément l'hétérogénéité de fixation du traceur dans le volume tumoral et donc établir un profil plus complet de la lésion.

L'analyse de texture des images TEP connaît un véritable engouement depuis 2009 et s'inscrit dans le concept de Radiomique. Certains auteurs ont montré que cette technique permettait de prédire la réponse aux traitements ou la survie du patient, à partir d'un examen TEP avant traitement, notamment pour des cancers pulmonaires, de l'œsophage, du col de l'utérus ou ORL. Toutefois, l'analyse de la littérature fait apparaître des protocoles de calcul variés et des conclusions parfois contradictoires entre les études.

Dans ce contexte, l'objectif de ce travail de thèse a été de développer et d'évaluer ces méthodes d'analyse de texture à partir des images TEP pour caractériser plus précisément l'activité métabolique des tumeurs.

Le premier chapitre de ce manuscrit rappelle le principe de l'imagerie TEP et décrit les méthodes d'analyse des images à partir d'index conventionnels ou en cours de développement. Nous présentons ensuite la place de cette technologie dans la prise en charge du patient en oncologie.

Le second chapitre expose la problématique de la caractérisation de l'hétérogénéité tumorale en cancérologie et le principe de l'analyse de texture des images TEP en détaillant le protocole de calcul ainsi que la définition des matrices et des index de texture. Nous présentons également un état de l'art portant sur cette thématique en analysant précisément l'influence des choix méthodologiques, la robustesse des paramètres et en comparant les conclusions obtenues dans les différentes études cliniques.

Le troisième chapitre détaille les résultats de l'étude méthodologique que nous avons menée pour identifier les paramètres influençant les index de texture comme la méthode de segmentation ou l'étape de discrétisation. En particulier, des recommandations sont données sur le protocole de calcul. Ce travail est complété par la réalisation de simulations afin de fournir des points de repère pour une interprétation en clinique de ces index.

À partir de l'étude de trois cohortes et en s'appuyant sur les recommandations établies précédemment, le quatrième chapitre décrit nos résultats concernant l'évaluation clinique des index de texture. Nous comparons la texture mesurée dans différents types de tissus et nous analysons les capacités des index à distinguer certaines caractéristiques histologiques des lésions.

Enfin, dans le cinquième chapitre, nous nous intéressons à une analyse multi-échelle de la caractérisation de l'hétérogénéité tumorale. Sur des modèles animaux, nous comparons la texture mesurée *in vivo* et *ex vivo* afin d'identifier le potentiel et les limites de l'analyse de texture des images TEP pour l'utiliser judicieusement en clinique.

En conclusion, nous faisons le bilan des résultats obtenus et dégageons les perspectives ouvertes par l'analyse de texture et le concept de radiomique.

Chapitre 1

La tomographie par émission de positons en oncologie

Ce chapitre présente brièvement le principe de la tomographie par émission de positons (TEP), des radiopharmaceutiques à la reconstruction des images. Nous exposons ensuite l'analyse des images TEP à partir des index classiques, utilisés en routine, et des paramètres en cours de développement en recherche. Enfin, nous détaillons la place de cet examen dans la prise en charge en oncologie.

1.1 Imagerie TEP en clinique

Dans cette première partie, nous allons aborder le processus de formation des images TEP, présenter les caractéristiques d'un examen et détailler les différents traceurs pouvant être utilisés en oncologie.

1.1.1 Principe de l'imagerie TEP

Les différentes modalités d'imagerie médicale peuvent être classées en deux grandes familles. La première regroupe les modalités d'imagerie anatomique qui permettent de visualiser la morphologie ou l'anatomie, c'est le cas de la radiologie X, la tomodensitométrie (TDM), l'imagerie par résonance magnétique (IRM) ou l'échographie. A l'opposé, la seconde regroupe les modalités d'imagerie fonctionnelle qui permettent d'étudier des phénomènes biochimiques ou physiologiques, grâce notamment à l'IRM fonctionnelle ou à la tomographie par émission de positons (TEP). Cette dernière repose sur l'injection au patient d'un traceur faiblement radioactif qui caractérise une fonction biologique que l'on souhaite visualiser et/ou mesurer *in vivo* et de façon non invasive. Le principe de l'imagerie TEP est de modifier la structure moléculaire du traceur, ciblant le processus biologique d'intérêt, en y intégrant un isotope radioactif. Les particules émises dans le patient à la suite de la désintégration de cet isotope sont détectées par un dispositif externe. Des algorithmes de reconstruction tomographique sont ensuite mis en œuvre pour cartographier en trois dimensions la répartition du traceur dans l'organisme du patient. La construction d'une image TEP repose donc sur trois éléments majeurs : le radiopharmaceutique, la détection et la reconstruction.

1.1.1.1 Radiopharmaceutiques

Pour être utilisé en imagerie TEP, le traceur doit présenter certaines caractéristiques [Phair, 1997] :

- une grande spécificité : il doit s'accumuler préférentiellement là où la fonction tracée est présente et avoir moins d'affinité là où elle ne l'est pas ;

- une cinétique compatible avec une acquisition clinique (durée d'immobilisation du patient et de l'examen) ;
- être l'analogue d'une molécule déjà présente dans l'organisme, le marquage ne devant pas modifier le comportement chimique étudié.

Dans le cas de l'imagerie TEP, le marqueur utilisé est un isotope émetteur β_+ . Cette émission se caractérise par une expulsion d'un positon et d'un neutrino suite à un excès de protons dans le noyau :



X représente le radioélément père, Z son nombre de protons, A son nombre de nucléons. Y est l'élément produit de la désintégration, ν_e le neutrino et β_+ la particule émise. De par la concentration très élevée d'électrons dans la matière, le positon émis parcourt une très faible distance avant de rentrer en collision avec un électron. Les deux antiparticules subissent une annihilation (Figure 1.1), ce qui conduit à la production de deux photons, de même énergie (511 keV) et qui s'éloignent l'un de l'autre dans deux sens opposés avec un angle d'émission de $180^\circ \pm 0,2^\circ$ [Bailey et al., 2005].

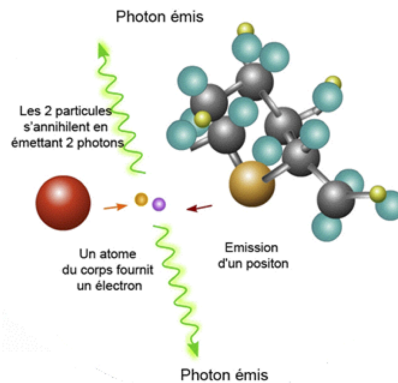


FIGURE 1.1 – Quand le positon rencontre un électron du milieu, une annihilation se produit et conduit à la formation de deux photons émis en sens opposés ($180^\circ \pm 0,2^\circ$). Source : CEA.

Les principaux émetteurs de positons utilisés en TEP sont : l'oxygène-15 (^{15}O), le carbone-11 (^{11}C) et le fluor-18 (^{18}F). Ces radio-isotopes peuvent être produits par bombardement de particules chargées sur une cible dans un accélérateur de particules (cyclotron). La demi-vie relativement courte de ces isotopes impose une production sur un site proche du centre d'imagerie. L'isotope le plus souvent utilisé est le ^{18}F ; il a une demi-vie de 110 minutes ce qui permet une élimination rapide, un temps limité d'exposition aux rayonnements pour le patient, mais nécessite des unités de production à proximité des utilisateurs.

1.1.1.2 Détection

L'objectif de l'imagerie TEP est de localiser le site d'émission des deux photons émis à 180° . Pour ce faire, un système de détection en coïncidence est utilisé, formé de cristaux disposés en couronne autour du patient, d'environ 80 à 90 cm de diamètre (Figure 1.2). Plusieurs dizaines de ces anneaux de détection sont ensuite associés en série pour former un tomographe TEP qui couvre un champ de vue axial d'environ 15 à 20 cm [Tarantola et al., 2003]. Ces anneaux sont constitués de cristaux scintillants suffisamment denses et épais afin d'arrêter majoritairement des photons de 511 keV. La droite joignant les deux détecteurs ayant collecté les deux photons en coïncidence est appelée ligne

de réponse.

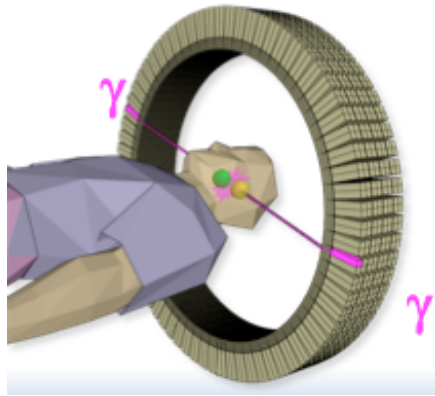


FIGURE 1.2 – Les cristaux sont organisés en couronne autour du patient pour recueillir les deux photons produits par l’annihilation. Source : Wikipedia.

Les photons interagissent majoritairement dans les cristaux par effet photoélectrique. Ils créent ainsi de multiples photons d’énergie plus faible, proportionnellement à la quantité d’énergie qu’ils déposent. Ces photons sont ensuite en partie collectés, le plus souvent, par des tubes photomultiplicateurs qui convertissent le flux de photons en signal électrique et l’amplifient. Les cristaux sont généralement regroupés en blocs (Figure 1.3) afin d’améliorer l’efficacité de détection, la localisation et le rendement. Pour la détection en coïncidence, un système discrimine les interactions qui ont lieu dans une fenêtre temporelle de l’ordre de 5 à 12 ns.

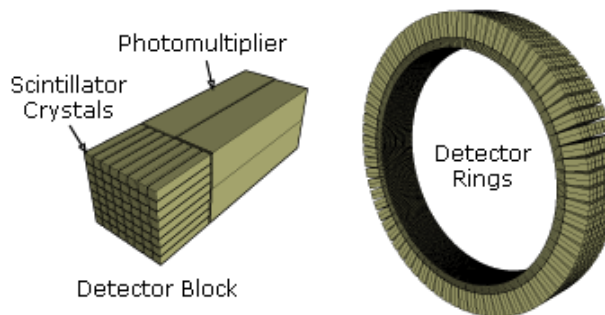


FIGURE 1.3 – Les cristaux sont regroupés en blocs. Source : Wikipedia.

Les imageurs TEP les plus récents mesurent le décalage entre les temps d’arrivées des deux photons sur les détecteurs (Figure 1.4), appelé aussi *temps de vol* (*time of flight* en anglais abrégé TOF). En effet, si l’annihilation a lieu à égale distance des deux cristaux, les photons vont être détectés simultanément. Au contraire, si l’annihilation est plus proche d’un des détecteurs, l’arrivée des deux photons va être décalée dans le temps. Ce décalage temporel fournit une information supplémentaire lors de la reconstruction et permet de déterminer plus précisément le lieu d’annihilation.

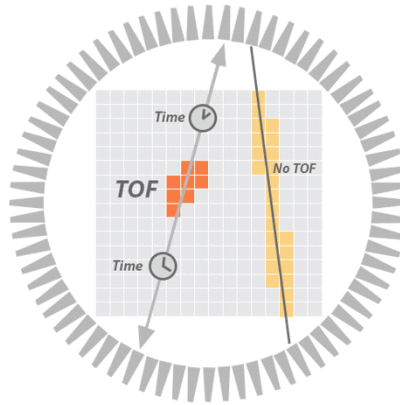


FIGURE 1.4 – Les systèmes TEP *temps de vol* ou *time of flight* (TOF) utilisent le décalage temporel d'arrivée des deux photons sur les détecteurs pour améliorer la localisation spatiale du lieu d'annihilation. Avec la technologie TOF, le lieu d'émission des photons est réduit à une faible portion de la ligne de réponse, à l'inverse des technologies non TOF. Source : Toshiba.

Le système de détection en coïncidence (Figure 1.5) recueille le signal de deux photons émis à la suite de l'annihilation d'un positon (coïncidence vraie) mais également des événements parasites qui ne proviennent pas du même positon. Ces coïncidences peuvent alors être de deux types :

- les coïncidences diffusées : elles correspondent à une coïncidence où au moins l'un des deux photons a subi une ou plusieurs diffusions dans le patient ou dans le système de détection avant d'atteindre le détecteur ;
- les coïncidences fortuites : elles correspondent à une coïncidence entre deux photons issus de deux annihilations distinctes.

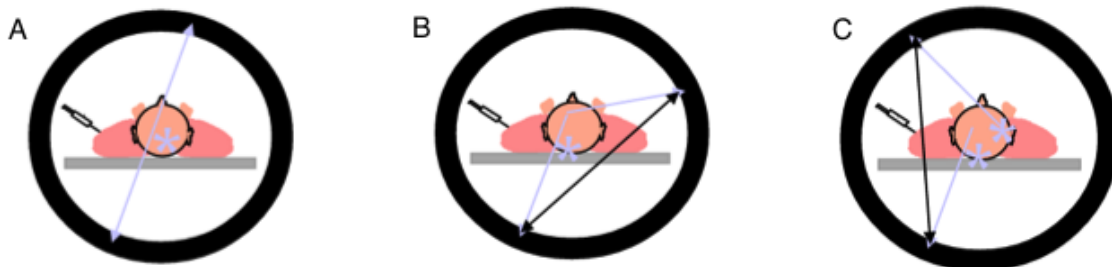


FIGURE 1.5 – Les systèmes de détection peuvent recueillir trois types de coïncidences : les coïncidences vraies (A), les coïncidences diffusées (B) et les coïncidences fortuites (C). Source : Irène Buvat.

1.1.1.3 Reconstruction

Pendant l'acquisition, les signaux détectés sont stockés sous forme de sinogrammes. Un sinogramme décrit les projections mesurées sur 180° autour du patient, en fonction de l'angle d'incidence, dans un tableau en deux dimensions comme présenté sur la Figure 1.6. Chaque ligne du sinogramme représente ainsi le nombre d'événements détectés dans toutes les lignes de réponse parallèles formant un même angle par rapport à l'axe du tomographe.

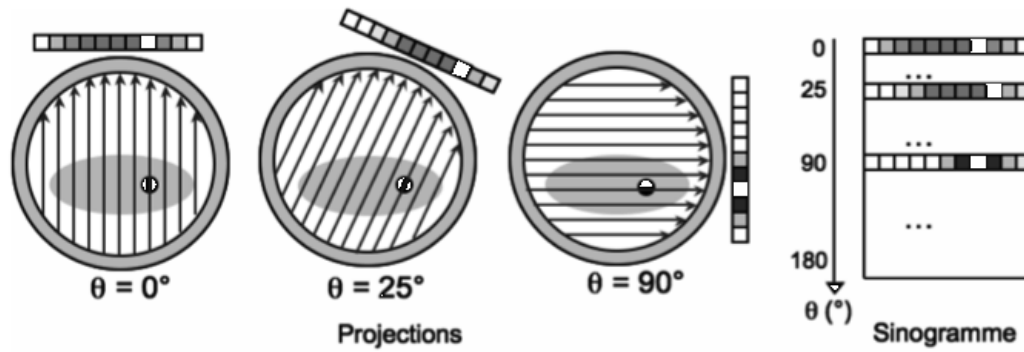


FIGURE 1.6 – Exemple de la formation de trois lignes du sinogramme suivant l’angle de projection.
Source : Perrine Tytski.

L’objectif de la reconstruction tomographique est de générer l’image tridimensionnelle de la distribution de radiopharmaceutique dans le patient à partir de ces sinogrammes. L’image reconstruite se présente alors comme une matrice de voxels. Deux grands types d’algorithmes de reconstruction sont utilisés : les méthodes analytiques et les méthodes itératives. Les méthodes analytiques, telles que la rétroprojection filtrée [Zeng, 2001], étant moins utilisées en routine clinique, nous nous intéresserons ici uniquement aux méthodes itératives.

Les algorithmes de reconstruction itératifs requièrent une estimation de l’opérateur de projection du système R , opérateur mathématique qui modélise la transformation entre l’objet f et ses projections p . Il s’agit donc de retrouver f dans cette équation connaissant l’opérateur R et les projections mesurées p :

$$p = R \times f \quad (1.2)$$

L’objectif est d’identifier, voxel par voxel, la distribution de l’activité du radiopharmaceutique la plus compatible avec les sinogrammes mesurés, sachant que l’inversion directe de la matrice de projection n’est pas réalisable. Pour retrouver f , on minimise itérativement une fonction de coût définie comme la distance entre les sinogrammes qui correspondent à la distribution d’activité estimée et les sinogrammes effectivement acquis. L’opérateur de projection R peut prendre en compte seulement la géométrie du détecteur mais également les effets physiques tels que la diffusion, l’atténuation ou la résolution spatiale limitée du détecteur.

En routine clinique, les algorithmes de reconstruction les plus utilisés reposent sur des approches statistiques et supposent notamment que le bruit affectant les données mesurées, suit une distribution de Poisson. Par exemple, les algorithmes appelés MLEM, *Maximum-Likelihood Expectation Maximization* [Shepp and Vardi, 1982], minimisent la fonction de coût en utilisant le maximum de vraisemblance. Une version accélérée, appelée OSEM pour *Ordered Subset Expectation Maximization*, minimise la fonction de coût, non plus sur toutes les projections, mais sur des sous-groupes répartis uniformément autour du volume de détection [Hudson and Larkin, 1994]. La qualité des images est alors déterminée par le nombre d’itérations et par le nombre de sous-groupes utilisés. La méthode RAMLA, *Row Action Maximum Likelihood Algorithm*, utilise quant à elle un nombre de sous-ensembles égal au nombre de projections [Browne and de Pierro, 1996]. La convergence de l’algorithme est contrôlée par un paramètre de relaxation. La méthode RAMLA peut également utiliser des volumes sphériques qui s’intersectent (BLOBS ou *spherically-symmetric basis functions*) à la place des voxels.

Les détecteurs TEP sont désormais systématiquement couplés à une machine TDM qui apporte une information anatomique aidant à l'interprétation des images et facilitant la correction de l'atténuation. En effet, les deux photons issus de l'annihilation peuvent interagir dans le milieu avant d'atteindre les détecteurs et réduire le nombre de coïncidences détectées. L'atténuation est ainsi responsable d'une sous-estimation d'activité supérieure à 70% [Buvat, 2007b]. Depuis 2010, les développements instrumentaux ont également permis l'avènement de systèmes hybrides couplant la TEP à un système IRM. Toutefois, la correction d'atténuation directement à partir des images IRM est encore en voie de développement et fait l'objet de nombreuses recherches [Hofmann et al., 2009].

1.1.2 Imagerie TEP au 18F-FDG

Nous allons maintenant nous intéresser aux acquisitions TEP les plus réalisées en oncologie en détaillant le métabolisme du traceur et le déroulement de l'examen.

1.1.2.1 Métabolisme du 18F-FDG

Les cellules tumorales ont un métabolisme glucidique très perturbé [Warburg, 1956]. Ce défaut est compensé par une concentration plus importante de récepteurs glucidiques à leur surface et une consommation accrue de glucose par rapport aux cellules dites normales [Smith, 1999]. En utilisant un analogue du glucose intégrant un isotope radioactif, il est possible de repérer les cellules grandes consommatrices de sucre. Le 2-desoxy-2-18F-fluoro-D-glucose, appelé 18F-FDG, est un sucre dont un des groupements hydroxyles a été remplacé par un élément fluor-18 (Figure 1.7). Il entre dans les cellules par la même voie que le glucose non marqué et y est phosphorylé, mais la substitution du groupement OH empêche le 18F-FDG d'être dégradé et de subir les étapes de la glycolyse. En conséquence, le 18F-FDG sous sa forme phosphorylée reste majoritairement piégé dans la cellule jusqu'à la désintégration du fluor-18 en oxygène-18. La molécule se lie alors avec un ion H^+ et peut ensuite être métabolisée par la cellule puis éliminée. Le 18F-FDG s'accumule dans la cellule, proportionnellement à sa consommation de glucose, tandis que le radiopharmaceutique non consommé est éliminé par voie urinaire.

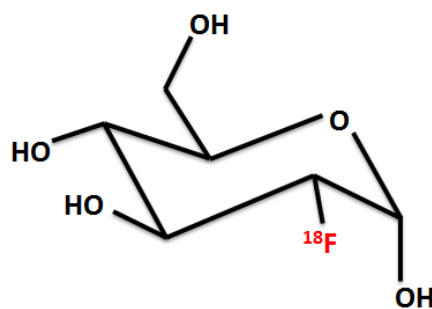


FIGURE 1.7 – Le 2-desoxy-2-18F-fluoro-D-glucose est un analogue du glucose et permet de visualiser et quantifier la consommation de glucose des cellules. Source : SeparationScience.

Le 18F-FDG va donc être assimilé en plus grande quantité dans les cellules tumorales que dans les cellules saines mais également dans les cellules soumises à des processus inflammatoires, les cellules cérébrales ainsi que les cellules musculaires sollicitées comme les cellules cardiaques ou dans les reins et la vessie compte tenu de l'élimination par voie urinaire (Figure 1.8).

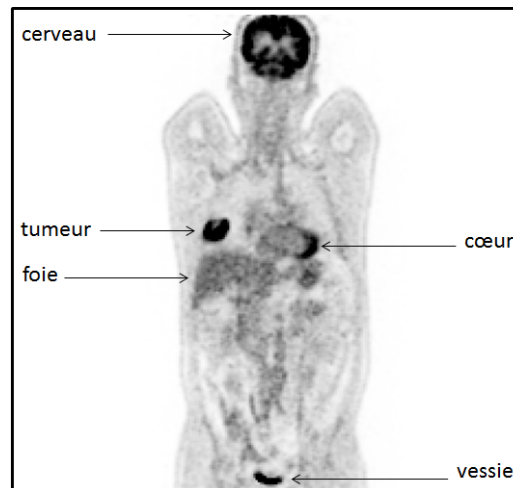


FIGURE 1.8 – Image TEP au 18F-FDG d'un patient atteint d'un cancer pulmonaire.

1.1.2.2 Déroulement d'un examen TEP au 18F-FDG

Possédant une autorisation de mise sur le marché (AMM) depuis 1998 en France, le 18F-FDG est le traceur le plus utilisé actuellement en TEP avec plus de 90% des examens. Pour maximiser la spécificité de l'examen, le patient est mis à jeun au moins 6 heures avant son admission dans le service de médecine nucléaire afin de limiter la compétition entre le glucose et le 18F-FDG dans les cellules. Le patient reçoit alors une injection de 18F-FDG par voie intraveineuse, de 3 à 3.5 MBq par kg. Il doit ensuite rester au repos pendant environ 60 minutes afin que le radiopharmaceutique se répartisse dans l'organisme. Après cette phase d'accumulation, le patient est placé dans le tomographe et doit rester immobile pendant environ 20 à 30 minutes.

La plupart des examens sont réalisés en mode statique, c'est-à-dire que les événements sont recueillis pour chaque pas de lit (environ 2 à 3 minutes). Pour les acquisitions dynamiques, les événements sont également recueillis par pas de lit mais avec plusieurs temps d'acquisition (de 5 secondes à plusieurs minutes). Ces acquisitions génèrent un volume de données beaucoup plus important et sont réservées aux protocoles de recherche et encore peu utilisées en clinique. Après l'acquisition, les images sont reconstruites puis interprétées par le médecin nucléaire.

1.1.3 Autres traceurs marqués au 18F en oncologie

En oncologie, d'autres traceurs ont été développés pour compléter les informations fournies par le 18F-FDG. Ils servent notamment à palier son manque de spécificité, lié à son accumulation dans des cellules non tumorales et son manque de sensibilité pour les tumeurs bien différenciées ou lentement évolutives.

1.1.3.1 Traceurs possédant une autorisation de mise sur le marché

Actuellement en clinique, trois traceurs marqués au fluor 18 possèdent une AMM depuis 2010 [Soussan et al., 2013], en plus du 18F-FDG :

- *la fluorocholine marquée au fluor 18 (18F-FCH)* : c'est un analogue de la choline qui est un constituant nécessaire à la synthèse de la membrane phospholipidique des cellules, plus abondant dans les cellules qui se multiplient comme les cellules cancéreuses. Ce radiophar-

maceutique est utilisé pour certains cancers au faible métabolisme glucidique pour lesquels le ^{18}F -FDG n'est pas performant comme le carcinome hépato-cellulaire, ou quand la fixation du ^{18}F -FDG par des organes voisins occulte la fixation du tissu d'intérêt (comme pour la vessie dans le cas du cancer de la prostate) ;

- *le fluorure de sodium marqué au fluor 18* (^{18}F -FNA) : c'est un analogue de l'ion fluorure présent dans la matrice osseuse. Il permet donc d'étudier les métastases osseuses, mais ne permet pas de visualiser les autres tissus cancéreux. Néanmoins, sa très grande sensibilité peut être utile lorsque la scintigraphie osseuse classique est douteuse sur une petite lésion peu active sur le squelette ;
- *le fluorodihydroxy-phénylalanine marqué au fluor 18* (^{18}F -LDOPA) : c'est un analogue de la phénylalanine qui intervient dans le transport actif membranaire et le stockage dans les granules de sécrétion de la dopamine. Le radiopharmaceutique permet donc d'étudier la synthèse de la dopamine pour le bilan des tumeurs endocrines ou cérébrales. Outre ces applications en oncologie, le ^{18}F -LDOPA peut notamment être utilisé en neurologie pour le diagnostic des syndromes parkinsoniens [Darcourt et al., 2014].

1.1.3.2 Traceurs en cours d'évaluation

Pour augmenter l'éventail des possibilités de l'examen TEP, d'autres traceurs sont actuellement à l'étude et sont basés sur les biocaractéristiques des tumeurs comme l'augmentation du métabolisme, l'hyper-prolifération, l'angiogenèse, l'hypoxie, l'apoptose ou encore sur des biomarqueurs spécifiques des cellules cancéreuses (antigènes ou récepteurs) [Zhu and Shim, 2011]. Les traceurs marqués au fluor 18 en cours d'évaluation sont résumés sur la Figure 1.9.

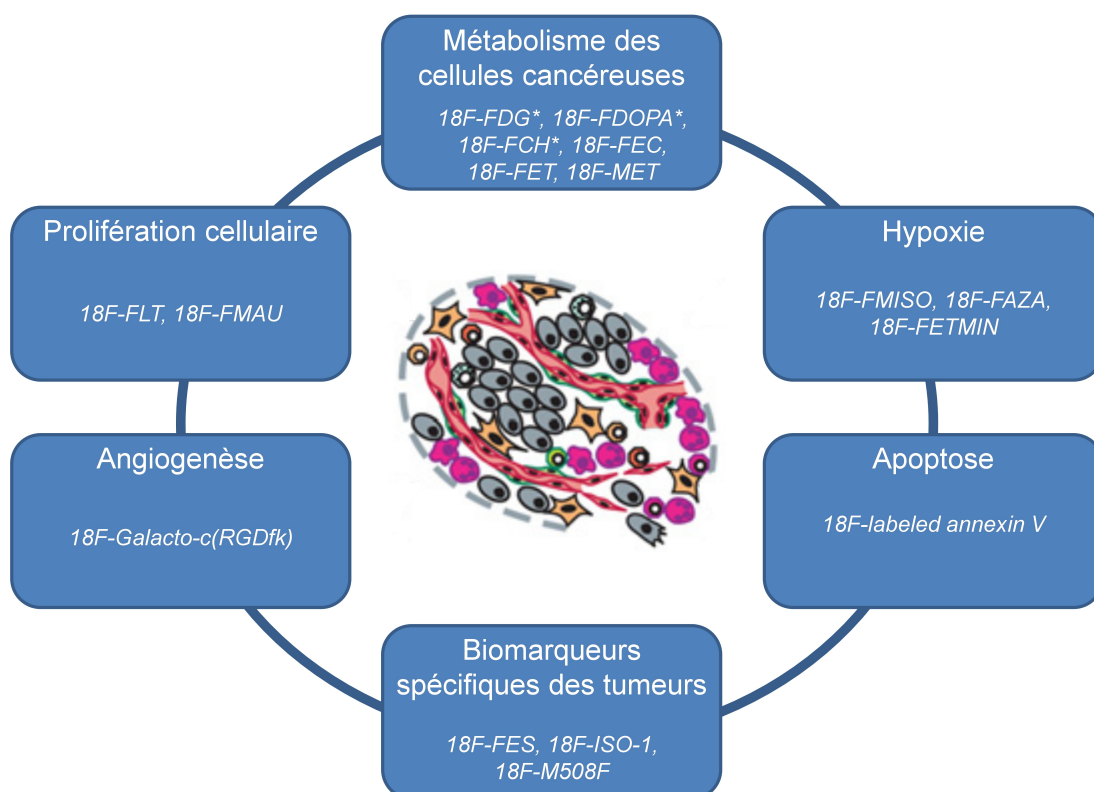


FIGURE 1.9 – Radiopharmaceutiques utilisés en clinique (*) ou en cours d'évaluation, marqués au fluor 18 suivant les biocaractéristiques tumorales qu'ils exploitent.

Après avoir présenté le processus de formation et de reconstruction des images TEP, nous allons maintenant décrire les méthodes qui permettent l'analyse des données.

1.2 Analyse des images TEP

La première étape de l'analyse des images TEP pour le médecin nucléaire consiste en une analyse visuelle lui permettant de repérer le nombre et la localisation des foyers cancéreux. Une analyse quantitative est ensuite réalisée pour mesurer le métabolisme des tumeurs. En oncologie, après avoir estimé l'activité métabolique et le volume des lésions, les médecins peuvent utiliser des index pour caractériser plus précisément l'activité tumorale.

1.2.1 Estimation de l'activité métabolique des tumeurs

1.2.1.1 Mesure du métabolisme du glucose

1.2.1.1.1 Le principe

Le modèle complet pour décrire l'évolution de la fixation du glucose est un modèle à trois compartiments (Figure 1.10). Il relie les concentrations de 18F-FDG dans les capillaires, dans le tissu et la concentration de 18F-FDG phosphorylé, piégé dans les cellules [Carson, 2005].

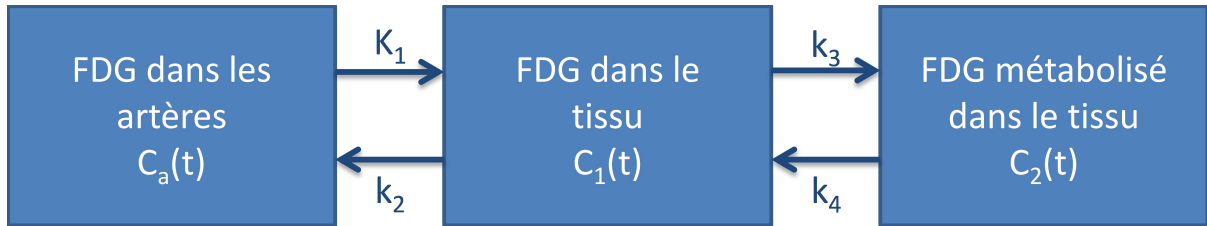


FIGURE 1.10 – Ce modèle cinétique à trois compartiments et quatre constantes d'échanges permet de décrire l'évolution de la fixation du glucose.

C_a correspond à la concentration d'activité de 18F-FDG dans les artères (kBq/mL), elle est appelée fonction d'entrée. C_1 et C_2 représentent respectivement les concentrations d'activité de 18F-FDG et de 18F-FDG métabolisé dans le tissu (kBq/mL). Les constantes d'échange K_1 ($mL/min/g$) et k_2 (min^{-1}) décrivent le transport du 18F-FDG vers et en dehors des membranes cellulaires ou des capillaires, k_3 représente le taux de phosphorylation du 18F-FDG et k_4 le taux de déphosphorylation. Le modèle permet donc de décrire l'évolution des concentrations dans les compartiments du 18F-FDG circulant et métabolisé :

$$\frac{dC_1}{dt} = K_1 C_a(t) - k_2 C_1(t) - k_3 C_1(t) + k_4 C_2(t) \quad (1.3)$$

$$\frac{dC_2}{dt} = k_3 C_1(t) - k_4 C_2(t) \quad (1.4)$$

Le plus souvent la constante de déphosphorylation k_4 est négligée puisque les tumeurs ont une activité phosphatase faible. Le modèle à trois constantes de temps permet d'écrire la concentration de 18F-FDG dans la région d'intérêt à un instant t :

$$C_2(t) = K_i^* \int_0^t C_a(\tau) d\tau + V_d C_a(t) \quad (1.5)$$

Où $K_i^* = \frac{K_1 \times k_3}{k_2 + k_3}$ correspond au débit net entrant de 18F-FDG en min^{-1} . V_d est la fraction du volume sanguin contenu dans la région d'intérêt (sans dimension).

Le premier terme représente la quantité de 18F-FDG phosphorylé piégé dans la tumeur. Le second terme rend compte du 18F-FDG non phosphorylé qui, par exemple, circule dans les vaisseaux irriguant la tumeur. Lors des acquisitions dynamiques, l'échantillonnage temporel permet de décrire l'évolution de l'activité dans la tumeur à l'aide de courbes « activité-temps » ($C_2(t)$). La fonction d'entrée ($C_a(t)$), quant à elle, est estimée à l'aide de prélèvements artériels réguliers. Connaissant $C_2(t)$ et $C_a(t)$, on peut alors estimer K_i^* et en déduire le taux de métabolisme du glucose Mr_{glu} ($\mu\text{mol.mL}^{-1}.\text{min}^{-1}$) d'après l'équation 1.6 :

$$Mr_{glu} = K_i^* \times \frac{C_{glu}}{LC} \quad (1.6)$$

Avec C_{glu} la concentration plasmatique de glucose et LC (« *lumped constant* »), une constante qui rend compte de la différence de transport et de phosphorylation existant entre le glucose et le 18F-FDG.

1.2.1.1.2 Analyse cinétique complète

La résolution du modèle cinétique est l'approche la plus complète pour décrire le métabolisme du glucose. Elle permet d'estimer rigoureusement les constantes de temps et ne fait que très peu d'approximations. À partir d'une séquence dynamique, les étapes de résolution du modèle sont les suivantes [Mankoff et al., 2006] :

- réalisation de l'acquisition des images avec prélèvements artériels simultanés ;
- délimitation des régions d'intérêt en 4 dimensions (3 dimensions spatiales et 1 dimension temporelle) et extraction des courbes « activités-temps » sur les images ;
- résolution du modèle compartimental à partir des courbes « activités-temps » et des prélèvements artériels pour la fonction d'entrée.

C_{glu} est déterminé par la mesure de la glycémie du patient. La difficulté majeure de cette méthode pour une utilisation en clinique consiste à effectuer des prélèvements artériels réguliers et à immobiliser le patient au cours de l'acquisition (une heure). Ces prélèvements constituent une opération délicate, posent des problèmes de radioprotection des personnels et présentent des risques pour le patient [Hall, 1971].

1.2.1.1.3 Méthode graphique de Patlak

La méthode de Patlak [Patlak and Blasberg, 1985] est une méthode moins contraignante que l'analyse cinétique complète. Elle requiert également des prélèvements sanguins mais est moins sensible à la qualité des prélèvements et les images peuvent être acquises avec une cadence moins élevée. La

méthode consiste à représenter la concentration de 18F-FDG dans la région d'intérêt ($C_2(t)$), divisée par la concentration sanguine de 18F-FDG ($C_a(t)$), en fonction de l'intégrale de cette dernière concentration entre l'administration du traceur et l'instant t , divisée par $C_a(t)$:

$$\frac{C_2(t)}{C_a(t)} = K_i^* \times \frac{\int_0^t C_a(\tau) d\tau}{C_a(t)} + V_d \quad (1.7)$$

Le terme de déphosphorylation (k_4) est également négligé. La pente de la courbe $C_2(t)/C_a(t)$ en fonction de l'intégrale $\frac{\int_0^t C_a(\tau) d\tau}{C_a(t)}$ permet de déterminer K_i^* et donc d'estimer Mr_{glu} (équation 1.6).

Cette approche est plus robuste à l'estimation de la fonction d'entrée que la méthode complète car l'ajustement linéaire est moins sensible au bruit que les méthodes de régression non linéaire [Hoekstra et al., 2002]. Néanmoins, elle ne permet pas de déterminer séparément les constantes de temps.

1.2.1.1.4 Analyse cinétique simplifiée

Les méthodes d'analyse cinétique simplifiée facilitent la résolution du modèle complet. La méthode SKM (*Simplified Kinetic Method*) [Hunter et al., 1996] nécessite seulement l'acquisition d'une image statique (entre 45 et 60 min post-injection) et un prélèvement veineux tardif (55 min post-injection). Le prélèvement permet d'ajuster la courbe d'activité sanguine à partir de courbes « témoins ». Néanmoins, cette méthode néglige le 18F-FDG non métabolisé.

Une autre approche nommée SKA-M (*Simplified Kinetic Analysis using Multiple time points*) [Sundaram et al., 2004] requiert un seul prélèvement sanguin pour estimer la fonction d'entrée mais détermine la fixation tumorale sur plusieurs images, permettant ainsi de prendre en compte le 18F-FDG non métabolisé.

Une généralisation de SKA appelée ESKA (*Extension of the Simplified Kinetic Analysis*) a été proposée pour optimiser le modèle de la fonction d'entrée ([Hapdey et al., 2011]). À partir de l'ensemble des fonctions d'entrée mesurées chez des patients témoins, une analyse en composantes principales est réalisée. Les deux premières composantes (C_{p-1}^* et C_{p-2}^*) sont retenues pour estimer la fonction d'entrée de chaque patient (C_p^*) en utilisant l'équation suivante :

$$C_p^*(t) = A_1 C_{p-1}^*(t) + A_2 C_{p-2}^*(t) \quad (1.8)$$

où A_1 et A_2 sont des constantes déterminées par deux prélèvements sanguins : un après l'injection du radiopharmaceutique et l'autre en fin d'examen. Cette méthode conduit à une meilleure estimation de K_i^* et est plus robuste que la méthode SKA à divers protocoles d'injection et d'acquisition.

1.2.1.2 Du taux de métabolisme du glucose au SUV

1.2.1.2.1 Définition du SUV

Même si elle est l'approche la plus complète pour décrire le métabolisme du glucose, la résolution du modèle cinétique, même simplifiée, est difficile à utiliser en clinique. À partir du modèle cinétique complet, quatre simplifications ont été proposées pour conduire à une mesure semi-quantitative appelée SUV (*Standardized Uptake Value*) [Woodard et al., 1975] ou valeur de fixation normalisée, l'index le plus communément utilisé pour caractériser la fixation du 18F-FDG en TEP [Buvat, 2007a] :

- première simplification : elle consiste à faire l'hypothèse qu'à l'instant t de l'examen, la fixation du 18F-FDG mesurée représente la quantité de 18F-FDG piégée dans la tumeur. Or la fixation augmente continûment au-delà des 60 minutes post-injection [Hamberg et al., 1994]. La mesure de SUV dépend donc de l'instant t où l'examen est réalisé ;
- seconde simplification : la fixation mesurée correspond au 18F-FDG métabolisé ; la part de 18F-FDG non métabolisé se situant dans les vaisseaux irriguant la tumeur est donc négligée ;
- troisième simplification : les tissus non-tumoraux consomment une quantité totale de 18F-FDG proportionnelle à la masse du patient. Cependant, tous les tissus ne consomment pas la même quantité de 18F-FDG ; c'est le cas par exemple des tissus graisseux qui consomment moins de 18F-FDG que les autres tissus ;
- quatrième simplification : la compétition entre glucose et 18F-FDG est négligée. En effet, si la concentration plasmatique en glucose est élevée, les fixations de 18F-FDG seront moindres par rapport à un patient à jeun.

La définition du SUV correspond à la fixation du radiopharmaceutique F (kBq/mL) dans le tissu d'intérêt normalisée par l'activité injectée et corrigée de la décroissance radioactive au temps de l'acquisition $A(t_{acq})$ (kBq) divisée par le volume V (mL) du patient. En pratique, on considère que la masse volumique du patient est égale à celle de l'eau ($1 g/mL$), la valeur de V est donc remplacée par la masse du patient en grammes :

$$SUV = \frac{F}{\frac{A(t_{acq})}{V}} \quad (1.9)$$

Le SUV est une quantité sans dimension. Si le traceur se répartissait uniformément dans tout l'organisme, le SUV serait identique en tout point et égal à un.

1.2.1.2.2 Paramètres influençant la mesure de la concentration de FDG

1.2.1.2.2.1 Influence des paramètres biologiques

De par la troisième simplification, le calcul du SUV nécessite de faire l'hypothèse d'une dilution uniforme du traceur dans l'organisme. Cette approximation grossière a pour conséquence de rendre la valeur du SUV dépendante de la composition du patient et en particulier de sa proportion de tissus adipeux. Pour pallier à ce problème, des auteurs ont introduit une normalisation en fonction de la masse maigre, SUV_{LBM} (*Lean Body Mass*) [Sugawara et al., 1999, Morgan and Bray, 1994] :

$$Femmes : LBM(kg) = 1,07 \times masse(kg) - 148 \times \left(\frac{masse(kg)}{taille(m)} \right)^2 \quad (1.10)$$

$$Hommes : LBM(kg) = 1,10 \times masse(kg) - 120 \times \left(\frac{masse(kg)}{taille(m)} \right)^2 \quad (1.11)$$

Une autre méthode a été proposée ; elle consiste à normaliser la fixation du traceur par la surface corporelle totale, SUV_{BSA} (*Body Surface Area*) [Kim et al., 1994] :

$$BSA(m^2) = 0,007184 \times masse(kg)^{0,425} \times taille(cm)^{0,725} \quad (1.12)$$

Dans ce cas, le SUV n'est plus sans dimension mais s'exprime en $m^2 \cdot mL^{-1}$.

1.2.1.2.2.2 Influence des paramètres d'acquisition

Les paramètres d'acquisition et de reconstruction des images influent sur les mesures de fixation. Des corrections sont donc nécessaires afin d'estimer le plus correctement possible la concentration de ¹⁸F-FDG dans la tumeur [Buvat, 2007a] :

- la correction de l'atténuation [Kinahan et al., 2003] : elle consiste à corriger l'atténuation des photons issus de l'annihilation en cartographiant les milieux atténuants dans le patient. En clinique, la carte d'atténuation est obtenue à partir de l'image TDM, dont les valeurs exprimées en Unités Hounsfield sont converties en coefficients d'atténuation à 511 keV au moyen d'une relation bilinéaire ;
- la correction de la diffusion [Zaidi and Koral, 2004] : elle est réalisée systématiquement, puisque plus de 50% des coïncidences détectées sont des coïncidences diffusées ;
- la correction du mouvement respiratoire [Pépin et al., 2014] : elle permet de compenser la sous-estimation de l'activité dans les images due au mouvement respiratoire. Les lésions les plus concernées se situent dans les poumons, le foie, les reins et le pancréas. Des méthodes encore peu utilisées en clinique permettent de synchroniser les acquisitions TEP à la respiration ;
- la correction de volume partiel [Erlandsson et al., 2012] : l'effet de volume partiel provient de la résolution spatiale limitée des images TEP et de leur échantillonnage. Il conduit à un étalement du signal dans les images et à un mélange des signaux dans chaque voxel. L'effet de volume partiel produit une sous-estimation de l'activité dans les structures hyperfixantes dont la taille est inférieure à environ trois fois la résolution spatiale ;
- l'optimisation de la reconstruction tomographique [Hutton, 2011] : l'algorithme de reconstruction choisi ainsi que les paramètres utilisés, comme le nombre d'itérations ou le filtrage, déterminent la résolution spatiale des images reconstruites et leur niveau de bruit, deux caractéristiques influençant l'intensité du signal mesuré.

Après avoir exposé la signification du SUV et les biais qui peuvent affecter son estimation, nous allons aborder l'évaluation du volume métabolique à partir des images TEP.

1.2.2 Estimation du volume métabolique

La segmentation vise à décomposer l'image en un nombre limité de régions d'intérêt (ROI). En général, la segmentation des zones tumorales sur les images TEP consiste à distinguer les régions tumorales métaboliquement actives des régions environnantes (le fond). La délinéation des hyperfixations pathologiques à partir des images TEP a plusieurs applications :

- mesurer le volume métaboliquement actif afin de caractériser le stade tumoral initial et de suivre son évolution au cours du traitement ;
- quantifier l'activité métabolique moyenne ou d'autres paramètres sur le volume tumoral global ;
- définir le volume tumoral pour l'intégrer dans le plan de traitement en radiothérapie.

La délinéation manuelle est une approche intuitive et encore largement appliquée en routine clinique. Cette méthode présente de nombreux inconvénients [Nestle et al., 2005, Nestle et al., 2007] :

- elle est longue et fastidieuse à mettre en place car la délinéation doit se faire coupe par coupe ;

— elle est peu reproductible puisqu'elle est intra et inter-opérateur dépendante.

La nécessité d'utiliser des méthodes automatiques ou semi-automatiques s'est donc imposée comme le souligne [Foster et al., 2014]. Cependant, la mise en place de méthodes de segmentation est particulièrement complexe du fait des caractéristiques des images TEP : résolution spatiale limitée, taille des voxels, effet de volume partiel, bruit dans l'image, hétérogénéité des tumeurs et mouvements du patient. Ainsi, il n'est pas possible de visualiser des bords nets entre l'hyperfixation d'intérêt et les tissus environnants. Par conséquent, des méthodes spécifiques ont dû être développées pour segmenter précisément les tumeurs à partir des images TEP [Vauclin et al., 2010].

1.2.2.1 Méthodes basées sur des seuils d'intensité

Les méthodes de segmentation par seuillage discriminent les tissus tumoraux des tissus environnants en fonction de leur intensité, sans prise en compte d'informations spatiales. La méthode utilisée fréquemment en routine consiste à déterminer un seuil fixe pour segmenter la lésion. Ce seuil correspond soit à une valeur fixe de SUV, souvent 2,5 unités SUV [Nestle et al., 2005], soit à un pourcentage de la valeur maximale dans la tumeur, entre 40% et 50% suivant les études. Toutefois, la définition d'un seuil optimal pour segmenter une tumeur sur les images TEP dépend fortement de la différence d'activité entre la tumeur et le fond, ainsi que de la taille de la lésion. Or, ces paramètres ne sont pas pris en compte dans les méthodes de seuillage fixe. Des techniques de détermination du seuil optimal par seuillage adaptatif ont donc été développées afin d'intégrer ces différences d'activité. Une de ces méthodes développée par [Nestle et al., 2005] consiste à définir un seuil optimal égal à :

$$Seuil = \alpha I_{70} + I_{fond} \quad (1.13)$$

où α correspond à un facteur de calibration déterminé expérimentalement sur des données fantômes ; I_{70} correspond à l'activité moyenne dans le contour défini par un seuil de 70% de la fixation maximale dans la zone d'intérêt ; I_{fond} correspond à l'intensité du fond, calculée à partir de la moyenne des voxels constituant une coque de deux voxels d'épaisseur, située à six voxels du bord du masque défini par le seuil de 70% précédent (on ne considère que les voxels d'intensité inférieure à 2,5 unités SUV pour ne pas inclure des zones d'hyperfixations voisines).

D'autres approches reposent sur des méthodes itératives qui utilisent le contraste, la taille ou l'activité moyenne dans la région segmentée pour définir un seuil [Jentzen et al., 2007, Drever et al., 2007]. La détermination du seuil optimal se fait itérativement jusqu'à convergence, c'est-à-dire quand le volume segmenté n'évolue plus entre deux seuils successifs. Ces méthodes nécessitent également une étape initiale d'étalonnage afin de définir la fonction qui donne le seuil optimal en fonction des paramètres considérés.

1.2.2.2 Méthodes basées sur des critères morphologiques

Pour segmenter la zone d'intérêt à partir des images TEP, il est également possible d'utiliser des critères morphologiques comme la détection des contours [Hsu et al., 2008]. Des méthodes dérivatives utilisant la dérivée première ou seconde peuvent être mises en œuvre pour détecter de manière simple les ruptures dans l'image. Les maxima obtenus sont ensuite détectés le plus souvent par la méthode de la ligne de partage des eaux comme dans [Beucher and Meyer, 1992] ou [Shepherd et al., 2012].

1.2.2.3 Méthodes de classification

Plusieurs méthodes utilisent les informations des voxels dans une zone contenant la tumeur afin de classer les voxels appartenant à la tumeur et les voxels appartenant au fond.

Le premier type de méthodes concerne les approches par agrégation : chaque voxel est associé à un vecteur de caractéristiques (sa valeur, écart-type de son voisinage, évolution au cours du temps ...), ils sont ensuite regroupés si leurs caractéristiques sont communes ou proches [MacQueen, 1967, Wong et al., 2002].

Un second type de méthodes regroupe les classifications paramétriques, qui utilisent l'histogramme d'un volume d'intérêt autour de la tumeur pour classer les voxels en trois classes : les voxels du fond, les voxels de la tumeur et les voxels intermédiaires [Aristophanous et al., 2007]. Ces derniers, à cause de l'effet de volume partiel, contiennent à la fois l'information du fond et de la tumeur. La discrimination des classes se fait par la résolution d'un modèle de mélange gaussien pour la distribution d'intensité des voxels.

Enfin, un troisième type de méthodes utilise cette fois la classification avec modélisation statistique du bruit. En effet dans les images TEP reconstruites, le bruit est corrélé, c'est-à-dire qu'il dépend de la valeur du voxel ou de ses voisins. Il a donc été proposé d'introduire un a priori spatial : un voxel a plus de chance d'appartenir à la même classe que ses voisins qu'à une autre classe. Pour modéliser les interactions entre voxels voisins, les champs de Markov peuvent être utilisés : l'hypothèse est alors que la valeur d'un voxels dépend uniquement des voxels situés dans son voisinage. Ainsi, la probabilité d'appartenance à une classe est conditionnée par la probabilité des voxels du voisinage [Chen et al., 2001]. Une autre approche nommée FLAB (*Fuzzy Locally Adaptive Bayesian*) [Hatt et al., 2009] combine à la fois les informations provenant du bruit des images reconstruites et du flou lié à la résolution spatiale. Les voxels sont alors associés à deux ou trois classes suivant l'hétérogénéité de la tumeur.

1.2.3 Caractérisation tumorale

À partir des images exprimées en concentration d'activité normalisée (SUV) et/ou des segmentations de tumeurs, il est possible de calculer différents index caractérisant la lésion que nous allons maintenant présenter.

1.2.3.1 Méthodes classiques

Trois index issus directement des images normalisées en unités SUV sont majoritairement utilisés dans les études cliniques (Figure 1.11). Le SUVmax correspond au maximum d'intensité relevé dans le volume d'intérêt (VOI) tandis que le SUVmean rend compte de l'intensité moyenne dans ce même VOI. Le SUVmax étant potentiellement affecté par le bruit présent dans les images TEP et du fait de sa mesure sur un unique voxel, un paramètre nommé SUVpeak a été introduit. Le SUVpeak correspond à la moyenne de fixation dans une sous-région du VOI. Néanmoins, la définition de cet index ne faisait pas initialement consensus dans la littérature [Vanderhoek et al., 2012] et variait par sa localisation dans la tumeur (centrée sur SUVmax ou placée dans la zone de fixation la plus élevée), sa forme ou sa taille. En suivant la définition de [Wahl et al., 2009], le SUVpeak correspond à la fixation moyenne dans une sphère de 1 mL de sorte que le signal soit maximal (donc pas nécessairement centrée sur le voxel d'intensité maximale).

L'activité métabolique de la tumeur est reflétée certes par la quantité de traceur dans la zone d'intérêt mesurée par les trois index précédents mais aussi par la taille de cette hyperfixation. Le volume métaboliquement actif (VM) fournit donc également une information sur le métabolisme tumoral, même s'il est très dépendant de la méthode de segmentation choisie. Une combinaison de ces informations a été introduite dès 1999 par [Larson et al., 1999] : c'est la glycolyse de la lésion ou TLG (*Total Lesion Glycolysis*) qui correspond au produit de SUVmean par VM ou à l'intégrale du SUV sur le VOI.

Le SUVmax reste toutefois l'index le plus populaire en routine clinique ; il serait mesuré dans plus de 90% des cas [Beyer et al., 2011].

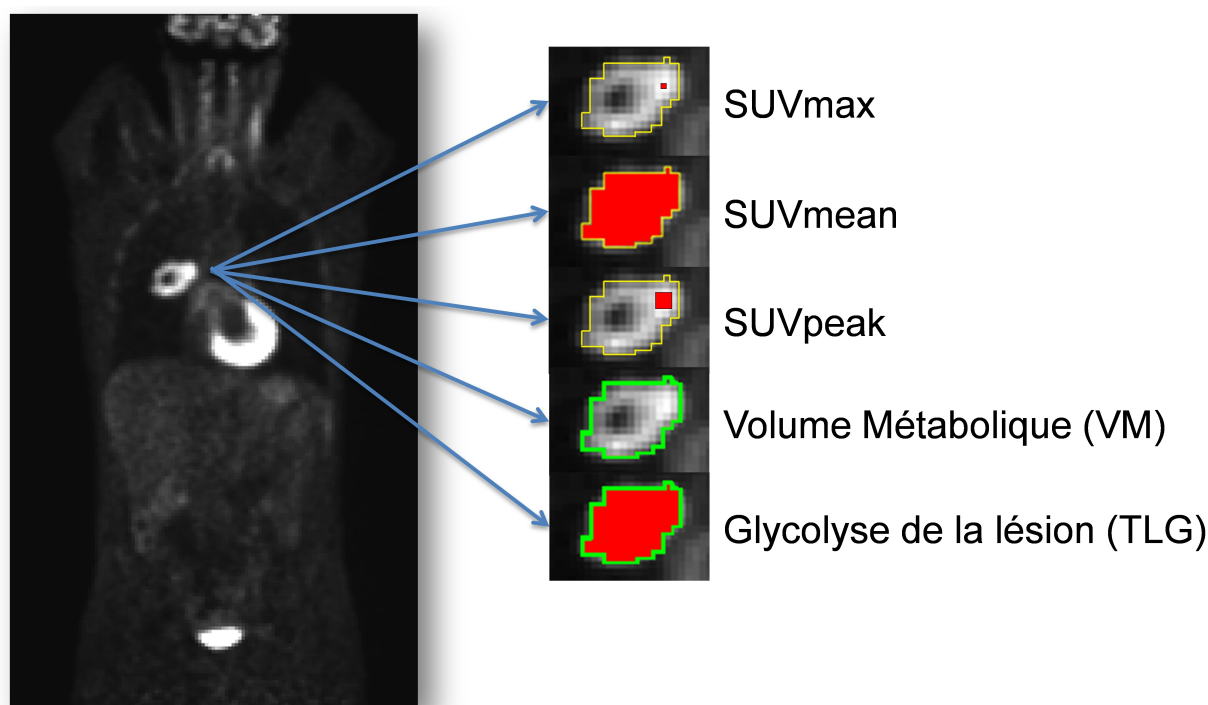


FIGURE 1.11 – Exemple de calcul des index classiques à partir de l'image d'une lésion segmentée.

1.2.3.2 Méthodes en développement

Au-delà des mesures « classiques » utilisées plus ou moins systématiquement en clinique et qui ont montré un intérêt notamment en oncologie, de nouveaux index ont été proposés pour rendre compte plus précisément des caractéristiques tumorales. Ces méthodes sont encore en cours de développement et de validation.

1.2.3.2.1 Index issus de l'histogramme

En complément des mesures « classiques » de SUV, il est possible de s'intéresser aux différences de niveaux de fixation des voxels présents dans le VOI. Pour ce faire, on peut calculer les statistiques du premier ordre qui sont issues de l'histogramme d'intensité. En plus du maximum et de la moyenne dans le VOI calculés via le SUVmax et le SUVmean, on peut extraire depuis l'histogramme les mesures répertoriées dans le Tableau 1.1.

Ces métriques décrivent des mesures globales sur la tumeur en ignorant les relations spatiales entre les voxels. Quelques études ont caractérisé leur intérêt en imagerie TEP. Dans le cas du cancer pulmonaire non à petites cellules, [Pyka et al., 2015] montrent que le COV n'est ni prédictif de la récurrence ni de la survie sur 45 patients, tout comme [Tixier et al., 2014b] sur 102 patients. Sur 27 patients atteints de cancers rectaux, le COV apparaît prédictif de la réponse et de la survie (Aire sous la courbe ROC, abrégée AUC, supérieure à 0,80) [Bundschuh et al., 2014] à l'opposé du skewness et du kurtosis. Enfin, sur une cohorte de 41 patients atteints de cancer de l'œsophage, ni l'écart-type, ni le skewness, ni le kurtosis n'apparaissent prédictifs de la réponse [Tixier et al., 2011].

Nom	Description	Formule
minimum	minimum d'intensité dans l'histogramme	SUV_{min}
écart-type	mesure de la dispersion des valeurs	$\sigma = \sqrt{\frac{1}{E} \sum_i (H(i) - SUV_{mean})^2}$
coefficient de variation	mesure de la dispersion relative	$COV = \frac{\sigma}{SUV_{mean}}$
coefficient d'aplatissement ou kurtosis	mesure l'aplatissement de la distribution (Figure 1.12).	$K = \frac{\frac{1}{E} \sum_i (H(i) - SUV_{mean})^4}{\left(\frac{1}{E} \sum_i (H(i) - SUV_{mean})^2 \right)^2}$
coefficient de dissymétrie ou skewness	mesure l'asymétrie de la distribution des valeurs d'intensité des voxels (Figure 1.12).	$S = \frac{\frac{1}{E} \sum_i (H(i) - SUV_{mean})^3}{\left(\sqrt{\frac{1}{E} \sum_i (H(i) - SUV_{mean})^2} \right)^3}$
énergie	mesure de l'uniformité de la distribution	$Energie_{Hist} = \sum_i p(i)^2$
entropie	mesure du désordre	$Entropie_{Hist} = - \sum_i p(i) \times \log(p(i))$

Tableau 1.1 – Résumé des index extraits de l'histogramme. E : nombre de voxels dans le VOI, $H(i)$: nombre de voxels d'intensité i , $p(i)$: probabilité d'apparition du niveau i .

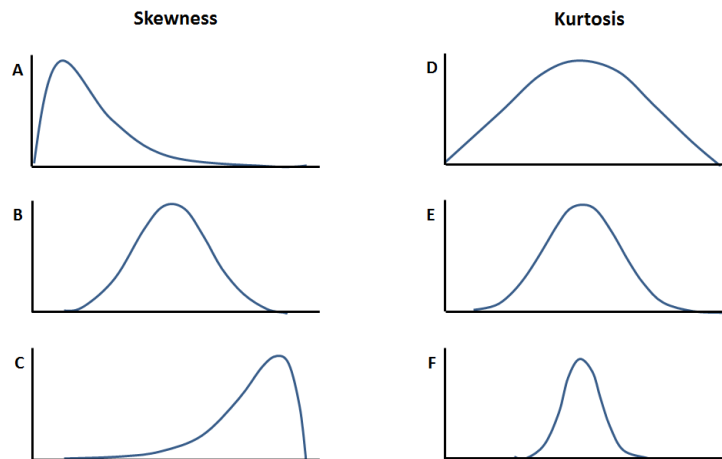


FIGURE 1.12 – Le *skewness* est positif (A) si la distribution est décalée à gauche, nul (B) pour une distribution symétrique et négatif (C) si la distribution est décalée à droite. Le *kurtosis* est inférieur à 3 (D) si les queues de distribution comptent plus d'observations que dans une distribution gaussienne, nul (E) si la distribution suit une loi gaussienne et supérieur à 3 (F) si les queues comptent moins d'observations.

1.2.3.2.2 Standardized added metabolic activity (SAM)

Le *Standardized added metabolic activity* (SAM) que l'on peut traduire par valeur de l'activité ajoutée normalisée, a été développé par [Mertens et al., 2012, Mertens et al., 2013]. Cet index mesure la glycolyse d'une lésion en corrigeant l'effet de volume partiel. Le SAM est calculé en traçant un premier VOI autour de la tumeur (VOI_1) et un second VOI plus grand (VOI_2) autour de VOI_1 . L'activité du fond est calculée en utilisant :

$$Activité_{fond} = \frac{TLG_{VOI_2} - TLG_{VOI_1}}{VM_{VOI_2} - VM_{VOI_1}} \quad (1.14)$$

On peut ensuite corriger la concentration d'activité issue de VOI_1 en soustrayant celle provenant du fond :

$$SAM = TLG_{VOI_1} - (Activité_{fond} \times VM_{VOI_1}) \quad (1.15)$$

Enfin, la valeur de SAM normalisé correspond au SAM divisé par l'activité du fond, car le fond et la tumeur sont affectés de la même façon par les paramètres d'acquisition :

$$SAM_{normalisé} = \frac{SAM}{Activité_{fond}} \quad (1.16)$$

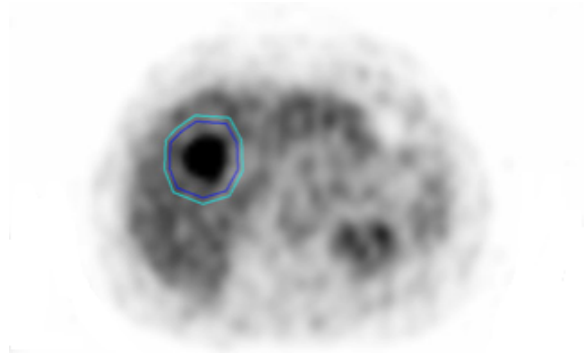


FIGURE 1.13 – Image TEP au 18F-FDG d'un patient présentant une métastase hépatique. VOI_1 représenté par le contour bleu et VOI_2 par le contour cyan, permettent de calculer l'index SAM. Source : [Mertens et al., 2012].

Le principal avantage de cette méthode est qu'elle ne nécessite pas une segmentation précise de la zone tumorale et qu'elle intègre une correction de l'effet de volume partiel si le contour de VOI_1 est suffisamment éloigné de la lésion.

Cet index a été étudié pour évaluer la réponse aux traitements et prédire la survie des patients atteints de métastases hépatiques [Mertens et al., 2012, Mertens et al., 2013] ou pour la caractérisation de cancers ORL [Kendi et al., 2015b, Kendi et al., 2015a].

1.2.3.2.3 Histogramme d'intensité cumulé

L'histogramme d'intensité cumulé correspond à la fraction du volume V_x ayant une fixation supérieure ou égale à $x\%$ de SUV_{max} en fonction de la fixation $I_{x\%SUV_{max}}$ (Figure 1.14). Dans la littérature il est nommé *cumulative SUV-volume histogram* (CSH) ou *intensity volume histogram* (IVH) et

se rapproche des histogrammes dose-volume utilisés en radiothérapie. À partir de cette courbe (Figure 1.14), il est possible d'extraire plusieurs index comme :

- la différence de fraction du volume entre I_{10} et I_{90} : $V_{10-90} = V_{10} - V_{90}$;
- l'aire sous l'histogramme nommée AUC-CSH ;
- la pente de la courbe.

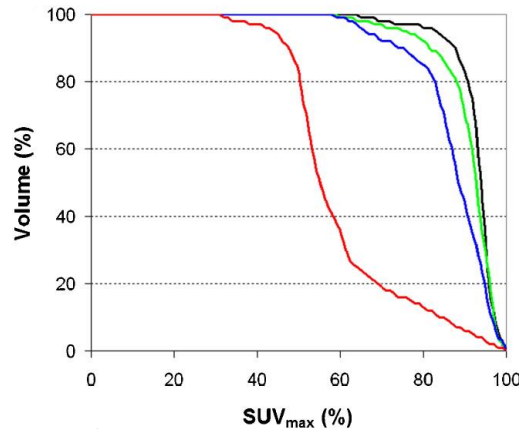


FIGURE 1.14 – Histogramme d'intensité cumulé représenté pour quatre modèles de tumeurs simulées. En rouge : tumeur hyperfixante au centre (SUV=12 au centre, 5 sur les bords). En bleu : tumeur hypofixante au centre (SUV=3 au centre, 5 sur les bords). En vert : tumeur entièrement hypofixante (SUV=3). En noir : tumeur moyennement fixante (SUV=5). Source : [van Velden et al., 2011].

Dans le cas du cancer de l'utérus, V_{10-90} conduit à une AUC de 0,65 contre 0,50 dans le cas des cancers ORL pour prédire la réponse du patient aux traitements [El Naqa et al., 2009]. La pente de l'histogramme serait corrélée à la récurrence post-radiothérapie pour des cancers pulmonaires non à petites cellules [Vaidya et al., 2012]. Van Velden et al. ont montré dans leurs travaux que AUC-CSH pouvait être considéré comme un index quantitatif de l'hétérogénéité de fixation du traceur [van Velden et al., 2011, van Velden et al., 2014]. Watabe et al. ont observé une différence de AUC-CSH entre les tumeurs gastro-intestinales et des lymphomes malins dans l'abdomen sur 21 patients [Watabe et al., 2011]. Dans le cas de cancers pulmonaires non à petites cellules, AUC-CSH est un facteur pronostique indépendant de la survie sans progression [Kang et al., 2014]. Par contre, dans les lymphomes diffus à grandes cellules B, le même index ne permet pas de différencier les patients répondeurs et non répondeurs, à la différence des SUV, du VM et du TLG [Tateishi et al., 2015]. Enfin, cette approche ne permet pas de distinguer des distributions spatiales différentes [Brooks, 2013].

1.2.3.2.4 Paramètres de forme

D'autres auteurs ont étudié, non plus la quantification de la fixation dans la tumeur, mais la forme de la lésion. Plusieurs index ont été utilisés :

- le nombre d'Euler : correspond au nombre d'objets connectés dans le VOI moins le nombre de trous ;
- la solidité : mesure la convexité du VOI. Il correspond à la proportion des voxels contenus dans l'enveloppe convexe contenant le VOI ;
- l'asphéricité : quantifie l'écart à une sphère parfaite (0 pour une sphère parfaite) $100 \times (\sqrt[3]{H} - 1)$ avec $H = \frac{Surface^3}{36\pi Volume^2}$.

El Naqa et al. [El Naqa et al., 2009] ont démontré que la solidité pouvait prédire la réponse aux traitements de patients atteints de cancer du col de l'utérus (AUC=0,72) ou ORL (AUC=0,85), tout comme le nombre d'Euler dans ce dernier cas (AUC=0,85).

Quant à l'asphéricité, elle a montré une valeur pronostique dans le cas des cancer ORL d'après [Apostolova et al., 2014b, Hofheinz et al., 2015] et des cancers pulmonaires [Apostolova et al., 2014a].

Une autre approche a été développée par [Eary et al., 2008] et consiste à mesurer l'hétérogénéité spatiale des mesures de SUV. Cette méthode modélise une distribution idéalisée de la fixation dans la tumeur suivant une ellipsoïde en trois dimensions. Le modèle considère que la fixation d'une tumeur décroît progressivement du centre vers les bords de l'ellipsoïde. L'index, nommé variable d'hétérogénéité ou HET, décrit le pourcentage de variabilité de la fixation des voxels qui n'est pas expliqué par le modèle ellipsoïde. Les auteurs ont montré que ce modèle apportait une information pronostique sur 238 patients atteints de sarcomes.

1.2.3.2.5 Index de texture

Depuis 2009 [Alic et al., 2014], la quantification de l'hétérogénéité tumorale grâce à une analyse de texture des images TEP a été suggérée par [El Naqa et al., 2009]. Ces paramètres seront détaillés précisément dans le chapitre suivant.

1.2.3.2.6 Analyse fractale

Afin de caractériser l'hétérogénéité spatiale des lésions, la mesure de la dimension fractale à partir des images TEP a été proposée. La géométrie fractale se définit comme la relation entre une mesure M et une échelle ϵ , exprimée comme :

$$M = k \epsilon^{-D} \quad (1.17)$$

avec k une constante d'échelle et D la dimension fractale. En appliquant différents seuils (M) successifs sur les images, on peut calculer le nombre de voxels (ϵ) ayant une fixation supérieure à ce seuil et tracer la courbe du logarithme népérien de M en fonction du logarithme népérien de ϵ . La pente de la courbe nous donne alors la dimension fractale (D) de l'image.

$$\ln\left(\frac{M}{k}\right) = -D \ln(\epsilon) \quad (1.18)$$

Miwa et al. [Miwa et al., 2014] ont ainsi montré que la mesure de D sur les images TEP permettait d'améliorer la différenciation entre les nodules pulmonaires malins et bénins comparé à SUVmax seul.

Après avoir présenté les méthodes d'analyse des images TEP utilisées en routine ou en cours de développement, nous décrivons l'utilisation en clinique des examens TEP au 18F-FDG.

1.3 Place de la TEP au 18F-FDG dans la prise en charge en oncologie

Depuis son développement en clinique, à la fin des années 90, l'imagerie TEP a pris une place de plus en plus importante dans la prise en charge du patient. En oncologie, cet examen peut intervenir à quatre étapes clés : lors du diagnostic, pour guider la radiothérapie, évaluer la réponse aux traitements et surveiller d'éventuelles récives.

1.3.1 Diagnostic et bilan d'extension

La TEP au 18F-FDG a comme première indication de statuer sur une anomalie « indéterminée » détectée par une autre méthode d'imagerie (échographie, radiographie, IRM) afin d'en préciser la nature (bénigne ou maligne). C'est notamment le cas lors de la découverte de nodules pulmonaires de plus de 10 mm : en l'absence d'hyperfixation sur l'examen TEP au 18F-FDG, il y a une très forte probabilité de bénignité [Soussan et al., 2013]. L'imagerie TEP permet également d'effectuer le bilan d'extension en recherchant des lésions métastatiques ou ganglionnaires qui pourraient modifier la stratégie thérapeutique initiale et qui ne seraient pas détectées par un bilan conventionnel. Enfin, elle peut être utilisée pour localiser une tumeur primitive en cas de maladie métastatique sans foyer primitif connu. Dans ces indications, l'analyse des images TEP est majoritairement visuelle, à la recherche du nombre et de la localisation des foyers tumoraux.

1.3.2 Planification de traitement en radiothérapie

En France, plus de la moitié des patients atteints d'un cancer reçoivent un traitement par radiothérapie. Cette méthode de traitement à base de rayonnements de haute énergie a pour objectif de détruire efficacement les cellules cancéreuses en préservant au maximum les tissus sains environnants. Cette irradiation des cellules tumorales entraîne des lésions au niveau de l'ADN qui empêchent leur prolifération et conduisent à la mort cellulaire.

Une étape indispensable avant le début du traitement consiste à délimiter les zones tumorales à irradier et les organes à risque pour lesquels l'irradiation peut causer des séquelles irréversibles. Dans la très grande majorité des cas, cette planification se fait à partir des images anatomiques TDM ou IRM. Trois volumes sont alors délimités (Figure 1.15) :

- le volume tumoral macroscopique (GTV : *Gross Tumour Volume*) : il correspond au volume macroscopique de la tumeur défini sur les images anatomiques ;
- le volume cible clinique (CTV : *Clinical Target Volume*) : il inclut en plus du volume précédent des marges pour englober les extensions tumorales possibles ;
- le volume cible final (PTV : *Planning Target Volume*) : il intègre à la fois le volume précédent ainsi que des marges de sécurité qui prennent en compte les mouvements du patient, le déplacement des organes et les erreurs de positionnement.

L'utilisation d'images fonctionnelles telles que les images TEP pourrait fournir davantage d'informations sur les caractéristiques biologiques des lésions et aider à une meilleure planification. Depuis 2009, l'Agence Internationale de l'Énergie Atomique [MacManus et al., 2009] recommande donc l'utilisation des images TEP en complément des images TDM :

- quand les lésions ne sont pas visibles sur les images TDM ou IRM ;
- pour éviter une irradiation des tissus environnants ne contenant pas de cellules tumorales ;

- quand une chimiothérapie complète le traitement par radiothérapie afin de mieux évaluer la réponse qu’avec un TDM seul ;
- pour la mise en place de thérapie qui s’adapte à la réponse tumorale, permettant ainsi une modification du volume cible pendant le traitement.

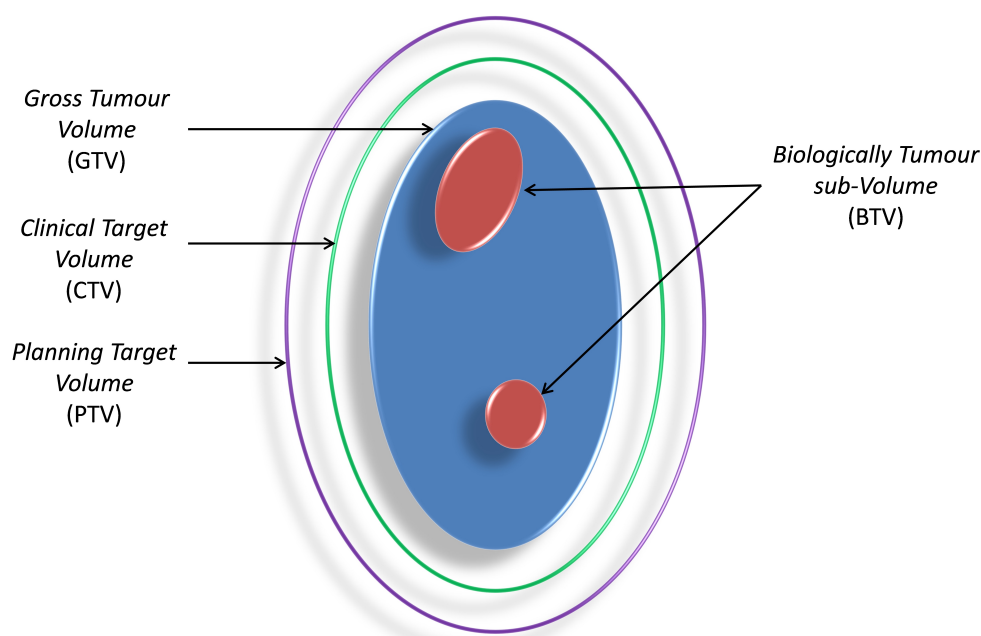


FIGURE 1.15 – Volumes d’intérêt pour la planification de traitement en radiothérapie.

L’utilisation des images TEP permettrait aussi d’administrer différentes doses d’irradiation à différentes sous-régions tumorales en définissant un quatrième volume d’intérêt (Figure 1.15), le BTV (*Biologically Tumour sub-Volume*). Il permettrait de sélectionner les zones radio-résistantes et donc d’augmenter localement la dose délivrée. Le contrôle tumoral local serait amélioré sans augmenter les effets secondaires [Aerts et al., 2009]. L’utilisation de traceurs autres que le 18F-FDG comme le 18F-FMISO ou le 18F-FAZA offrirait également aux médecins la possibilité de localiser les zones d’hypoxie qui contribuent à la résistance aux traitements notamment dans les cancers ORL et pulmonaires [Zips et al., 2012, Bollineni et al., 2014].

Cependant, le principal obstacle à l’utilisation de l’imagerie TEP repose sur la définition et l’obtention d’un volume métaboliquement actif qui soit fiable, reproductible et non-opérateur dépendant [Doll et al., 2014]. Or, à ce jour, aucune méthode de segmentation automatique ou semi-automatique ne fait consensus.

1.3.3 Évaluation de la réponse aux traitements

Lors de l’administration d’un traitement, l’objectif pour le médecin est de déterminer si la ou les lésions répondent à la thérapie partiellement, totalement, si elles sont stables ou en progression entre deux examens. Les conclusions de cette évaluation conditionnent soit la poursuite du traitement, soit son arrêt, soit sa modification afin d’optimiser la stratégie thérapeutique, tout en limitant les effets secondaires d’un traitement inefficace.

Dans un très grand nombre de cas, le suivi des patients se fait par des examens TDM. La réponse tumorale est donc évaluée majoritairement par des mesures de la taille des lésions sur des images anatomiques. Des critères ont ainsi été développés pour guider l'évaluation tumorale. Le premier a été présenté par l'Organisation Mondiale de la Santé (OMS) [WHO, 1979] et le second nommé RECIST (*Response Evaluation Criteria in Solid Tumors*) a été formulé en 2000 [Therasse et al., 2000] et mis à jour en 2009 [Eisenhauer et al., 2009]. Ils utilisent l'évolution des mesures anatomiques entre deux examens distants d'au moins 4 semaines. Une comparaison de ces deux méthodes est présentée dans le Tableau 1.2.

	OMS	RECIST 1.0	RECIST 1.1
Nombre de lésions	Toutes les lésions	10 au total avec 5 max/organe	5 au total avec 2 max/organe
Figure de mérite	Produit des plus grandes longueurs orthogonales (ΔP)	Mesure du plus grand axe sommée sur les lésions (ΔL)	Mesure du plus grand axe sommée sur les lésions (ΔL)
Réponse complète	Disparition de toutes les lésions	Disparition de toutes les lésions	Disparition de toutes les lésions et petit axe des ganglions lymphatiques pathologiques < 10 mm
Réponse partielle	$\Delta P \leq -50\%$ et ni apparition de lésions, ni progression	$\Delta L \leq -30\%$	$\Delta L(r) \leq -30\%$
Maladie stable	$-50\% < \Delta P < 25\%$	$-30\% < \Delta L < 20\%$	$-30\% < \Delta L(r) < 20\%$ et $\Delta L(a) < 5mm$
Maladie en progression	$\Delta P \geq 25\%$ ou apparition de nouvelles lésions	$\Delta L \geq 20\%$ ou apparition de nouvelles lésions	$\Delta L(r) \geq 20\%$ ou $\Delta L(a) \geq 5mm$ ou apparition de nouvelles lésions

Tableau 1.2 – Comparaison des trois critères anatomiques : OMS, RECIST 1.0 et 1.1. $\Delta L(r)$ représente l'évolution relative du plus grand axe et $\Delta L(a)$ l'évolution absolue.

Toutefois, les mesures anatomiques effectuées souffrent d'une variabilité inter et intra-opérateur et les critères de décision restent sujets à débat. De plus, les changements anatomiques évalués sont bien des marqueurs de la réponse tumorale, mais ces changements apparaissent postérieurement aux changements métaboliques. En effet, dans un premier temps, le processus de mort cellulaire n'induit pas de rétraction mais une modification de l'activité glycolytique. Les modifications fonctionnelles, quantifiables sur les images TEP apparaîtront donc préalablement aux modifications anatomiques.

Dès 1999, comme pour les images anatomiques, des critères ont été proposés pour évaluer la réponse aux traitements directement sur les images TEP. Les critères de l'EORTC (*European Organization for Research and Treatment of Cancer*) [Young et al., 1999] ont commencé par définir des recommandations pour la préparation du patient et la réalisation des examens TEP :

- le patient doit être à jeun et sa glycémie contrôlée ;
- l'acquisition des images doit être réalisée 50 à 70 min après l'injection du radiopharmaceutique et ce temps de référence doit être conservé pour tous les examens d'un même patient ;
- l'examen TEP pré-traitement doit être réalisé au maximum deux semaines avant le début de la thérapie ;

- un intervalle de 1 à 2 semaines est recommandé entre la fin du traitement et l'examen TEP post-traitement ;
- l'utilisation du SUV normalisé par la surface corporelle (SUV_{BSA}) est conseillée ;
- le calcul de SUVmax et de SUVmean sont recommandés.

Ces recommandations ont toutefois été établies seulement sur 10 études et pour 5 types de cancer. Elles ont néanmoins eu le mérite de poser les premières bases pour une standardisation de l'évaluation des traitements à partir des images TEP. En 2009, les critères PERCIST (*Positron Emission tomography Response Criteria In Solid Tumors*) [Wahl et al., 2009] ont été définis à partir de l'étude de plus de 300 publications et complètent ou modifient les recommandations précédentes par :

- un écart maximal de 15 min par rapport à l'examen de référence, entre le temps d'injection et le début de l'acquisition pour les examens de suivi ;
- la réalisation de toutes les acquisitions sur la même machine TEP ;
- l'injectio de la même dose de radiopharmaceutique ($\pm 20\%$) ;
- la possibilité de réaliser un examen TEP précoce (après le premier cycle de chimiothérapie) ;
- un intervalle de 10 jours entre la fin du traitement et l'examen TEP post-traitement ou 8 à 12 semaines après la radiothérapie ;
- l'utilisation du SUV normalisé par la masse maigre (SUV_{LBM}) ;
- la mesure de SUVpeak (moyenne maximale dans une sphère de volume 1 mL).

Les critères EORTC et PERCIST sont comparés dans le Tableau 1.3.

	EORTC	PERCIST
Réponse complète	Fixation égale à celle des tissus environnants	Fixation \leq fixation du compartiment sanguin et pas de nouvelles lésions
Réponse partielle	$\Delta SUV < -15/25\%$ après 1 cure ou $\Delta SUV < -25\%$ après plus qu'une cure	$\Delta SUV(r) < -30\%$ et $\Delta SUV(a) < 0,8$ et pas de nouvelles lésions
Maladie stable	$-15\% < \Delta SUV < 25\%$ sans augmentation visible du volume ($< 20\%$ du plus grand axe)	Ni réponse complète, ni réponse partielle, ni maladie en progression
Maladie en progression	$\Delta SUV > 25\%$ ou augmentation visible du volume ($> 20\%$ du plus grand axe) ou apparition de nouvelles lésions	$\Delta SUV(r) > +30\%$ ou $\Delta SUV(a) > 0,8$ ou augmentation du volume métabolique ou apparition de nouvelles lésions

Tableau 1.3 – Comparaison des critères d'évaluation de la réponse à partir des images TEP. $\Delta SUV(r)$ représente l'évolution relative du SUV et $\Delta SUV(a)$ l'évolution absolue.

Ces critères sont à l'heure actuelle utilisés dans les protocoles de recherche clinique (phases I à III) mais très peu en routine. Leur généralisation nécessite davantage de validations et une mise à disposition plus importante des examens TEP pour permettre à tous les patients de bénéficier d'un examen TEP pré-traitement et d'un autre au cours de la thérapie ou post-traitement.

1.3.4 Surveillance des récidives

Après traitement, l'imagerie TEP permet de statuer sur des anomalies radiologiques résiduelles en différenciant les lésions tumorales viables des modifications post-thérapeutiques (inflammation, fibrose ou nécrose). L'examen est également indiqué en cas de suspicion de récidives locales ou métastatiques.

Dans ce chapitre, nous avons brièvement présenté les différentes étapes qui permettent d'acquérir et de reconstruire les images TEP, et détaillé les méthodes d'analyses utilisées en pratique clinique ainsi que celles développées en recherche. L'imagerie TEP au 18F-FDG, par sa capacité à cartographier les foyers tumoraux et à quantifier leur métabolisme glucidique, en fait un examen de choix en oncologie et est amenée à se développer encore davantage dans les années à venir avec l'essor de la médecine personnalisée. Afin d'améliorer la caractérisation tumorale à partir de ces images, la mesure de l'hétérogénéité tumorale par une analyse de texture connaît un véritable engouement. Ces index et leur utilisation font l'objet d'une étude exhaustive dans le chapitre suivant.

Chapitre 2

Caractérisation de l'hétérogénéité tumorale

L'enjeu majeur de la cancérologie pour les années à venir est de proposer pour chaque patient un traitement personnalisé afin d'optimiser la réponse à la thérapie tout en minimisant les effets secondaires d'un traitement inefficace. Pour mettre en œuvre cette médecine personnalisée, l'un des défis est d'identifier très précocement les caractéristiques des lésions. Ainsi, l'estimation de paramètres de texture directement sur les images TEP a été proposée afin de quantifier l'hétérogénéité tumorale. Ce chapitre décrit ces paramètres, présente les principaux résultats de la littérature et détaille les problèmes qui se posent actuellement pour un transfert en clinique.

2.1 Motivations

2.1.1 Oncogenèse

Le terme cancer regroupe un ensemble de maladies pour lesquelles des cellules anormales se divisent sans contrôle et peuvent envahir les tissus voisins. Les cellules cancéreuses peuvent aussi se propager à d'autres parties du corps à travers les systèmes sanguins et lymphatiques (Source : NIH-National Cancer Institute). Il existe plusieurs types de cancers :

- le carcinome : cancer qui se développe dans la peau ou dans les tissus qui tapissent ou recouvrent les organes internes ;
- le sarcome : cancer qui débute dans les os, le cartilage, la graisse, le muscle, les vaisseaux sanguins ou tout autre tissu conjonctif ou de soutien ;
- la leucémie : cancer qui se développe dans les tissus hématopoïétiques, tels que la moelle osseuse et provoque un grand nombre de cellules sanguines anormales ;
- le lymphome et le myélome multiple : cancer qui débute par les cellules du système immunitaire ;
- les cancers du système nerveux central qui sont des cancers qui se développent dans les tissus cérébraux et de la moelle épinière.

Dans une revue de 2000 [[Hanahan and Weinberg, 2000](#)] et de 2011 [[Hanahan and Weinberg, 2011](#)], Hanahan et Weinberg ont pu identifier les dix mécanismes responsables de l'oncogenèse (cf. Figure 2.1) :

- le maintien des signaux de prolifération ;
- l'insensibilité aux signaux suppresseurs de croissance ;
- le dérèglement des métabolismes énergétiques ;
- la résistance à la mort cellulaire (apoptose) ;
- l'instabilité génomique et l'apparition de mutations ;
- la capacité d'angiogenèse ;

- la capacité d'invasion et de métastases ;
- l'inflammation favorisant le processus tumoral ;
- la prolifération illimitée ;
- l'échappement au système immunitaire.



FIGURE 2.1 – Mécanismes responsables de l'oncogénèse. Source : [Hanahan and Weinberg, 2011].

Ces caractéristiques conduisent à la conversion de cellules normales en cellules cancéreuses et à la formation d'une tumeur.

En 2011, 365 500 nouveaux cas de cancer ont été diagnostiqués en France. Le cancer touche un homme sur deux et une femme sur trois avant 85 ans avec un taux de guérison moyen de 60%. Les quatre principaux types de cancer sont les cancers de la prostate (71 000 nouveaux cas par an), les cancers mammaires (53 000), les cancers colorectaux (40 500) et les cancers du poumon (39 500). Ils représentent la première cause de décès prématuré avant 65 ans (Source : La ligue contre le cancer).

2.1.2 Hétérogénéité tumorale

Une tumeur n'est pas une entité homogène, constituée d'une seule sorte de cellules cancéreuses. Au contraire, le processus de mutation et de sélection donne naissance à des populations de cellules génétiquement distinctes. L'apparition de cette hétérogénéité tumorale est sujette à débat dans la communauté scientifique et deux théories cohabitent [Billaud, 2012]. La première se réfère à l'existence de cellules souches cancéreuses qui donneraient naissance à différents types de cellules et seraient responsables de l'hétérogénéité phénotypique d'une lésion [Shackleton et al., 2009]. La seconde se base sur un modèle darwinien de développement où la succession de mutations donne un avantage sélectif à une minorité qui se développe et coexiste au sein d'un même tumeur [Nowell, 1976].

Quatre types d'hétérogénéité tumorale peuvent être identifiés (cf. Figure 2.2) :

- l'hétérogénéité intra-tumorale qui représente les différences au sein des cellules d'une même tumeur ;
- l'hétérogénéité intra-métastase qui rend compte des différences au sein d'une seule métastase ;
- l'hétérogénéité inter-métastase qui représente les différences entre plusieurs sites métastatiques ;
- l'hétérogénéité inter-patient qui traduit les différences entre les tumeurs de plusieurs patients.

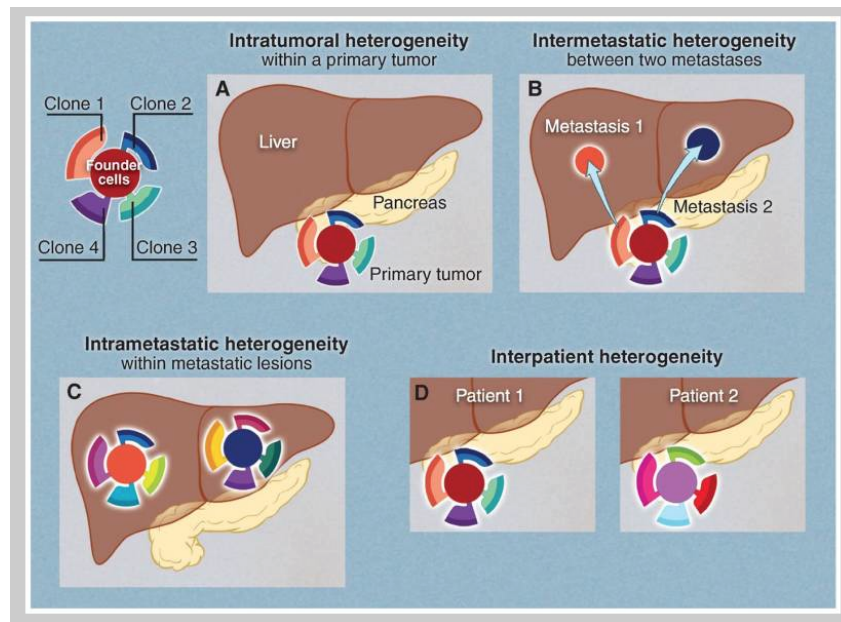


FIGURE 2.2 – Illustration des quatre types d'hétérogénéité tumorale possibles. Source : [Vogelstein et al., 2013].

Récemment, une équipe [Gerlinger et al., 2012] a testé si une seule biopsie permettait de repérer les altérations génétiques d'une tumeur à partir des biopsies multiples de 4 patients atteints d'un cancer rénal. Ils en ont conclu que 66 à 69% des mutations ne sont pas identifiées si l'analyse ne porte que sur un unique échantillon. Ils ont également démontré que la formation de ce type de cancer était compatible avec un processus de branchement évolutif qui se traduit par la génération de populations de cellules malignes qui occupent des régions distinctes de la tumeur dont une serait à l'origine des métastases.

La définition même de l'hétérogénéité intra-tumorale regroupe plusieurs notions suivant l'échelle d'observation, ce qui illustre la complexité biologique des tumeurs [O'Connor et al., 2015] :

- au niveau cellulaire, l'hétérogénéité peut être génomique, moléculaire ou phénotypique ;
- au niveau tissulaire, elle peut se refléter par des variations locales en termes d'architecture stromale, d'angiogénèse, de consommation d'oxygène ou de métabolisme glucidique ;
- au niveau macroscopique, l'hétérogénéité peut se traduire par des nécroses ou des infiltrations dans la tumeur.

La caractérisation de l'hétérogénéité tumorale est donc un véritable défi en oncologie comme le souligne [Marusyk and Polyak, 2010]. La coexistence de différents clones dans une tumeur influence l'évolution du cancer et doit être prise en compte dans la stratégie thérapeutique à adopter. C'est l'objectif de la médecine personnalisée : proposer un traitement optimal pour chaque patient suivant les caractéristiques propres de son cancer.

2.1.3 Nécessité d'identifier des biomarqueurs

Le développement de la médecine personnalisée en oncologie nécessite d'identifier des index fiables, appelés biomarqueurs, pour quantifier cette hétérogénéité à partir des données génomiques, protéomiques mais aussi des données d'imagerie. En effet, la réalisation de biopsies offre une très bonne résolution spatiale et permet d'effectuer de nombreuses analyses biologiques mais elles sont difficiles à réaliser, partielles, sujettes à erreurs et difficilement répétables au cours du traitement. Les systèmes d'imagerie quant à eux fournissent des informations sur la lésion entière et ses éventuelles métastases, de façon non-invasive et à plusieurs reprises au cours de la thérapie.

C'est ainsi que de nombreux groupes s'intéressent à l'analyse de texture notamment à partir des images TEP afin de quantifier l'hétérogénéité tumorale et de permettre une meilleure prise en charge du patient grâce à une analyse plus complète des examens déjà acquis en routine clinique.

2.2 Principe de l'analyse de texture

2.2.1 Protocole de calcul

Avant d'extraire les paramètres de texture, un protocole strict doit être respecté (Figure 2.3). Une première étape consiste à délimiter le VOI, soit par un contourage manuel, soit par des méthodes automatiques ou semi-automatiques comme présenté au Chapitre 1. Dans un second temps, une étape de discrétisation des intensités des voxels présents dans le VOI est nécessaire afin de réduire le bruit dans les images, de limiter la taille des matrices à calculer et de regrouper des voxels ayant des intensités proches. Cette étape sera plus largement détaillée au Chapitre 3. Ensuite, quatre matrices de texture peuvent être calculées, desquelles seront extraites plus d'une trentaine d'index de texture.

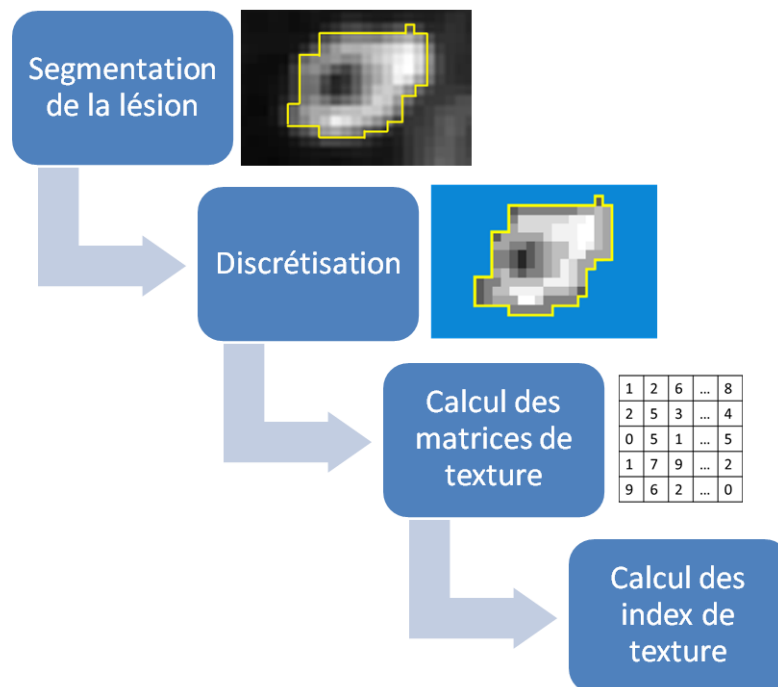


FIGURE 2.3 – Protocole de calcul des index de texture. Après avoir segmenté la lésion et discrétisé la valeur des voxels présents dans le VOI, les matrices de texture peuvent être calculées et les index de texture extraits.

2.2.2 Définition des matrices de texture

Dans cette partie, nous détaillons les quatre matrices de texture utilisées en imagerie TEP pour quantifier l'hétérogénéité tumorale : la matrice de co-occurrence, la matrice de différence de niveaux de gris et les matrices des longueurs des séries ou des zones homogènes. Pour faciliter la comparaison avec les résultats présentés dans la littérature et éviter des traductions approximatives, les noms des matrices et des index de texture ne seront pas traduits de l'anglais.

2.2.2.1 Matrice de co-occurrence

2.2.2.1.1 Calcul de la matrice

Proposée par [Haralick et al., 1973] dans les années 1970, la matrice de co-occurrence permet de réaliser une analyse statistique du second ordre. Cette matrice, largement utilisée en traitement d'images, permet de quantifier les relations entre des paires de voxels distant de d voxels dans la direction θ . La matrice résultante est une matrice carrée de taille $G \times G$, avec G le nombre de niveaux de gris dans l'image après discrétisation. Un élément (i, j) de la matrice de co-occurrence C se définit comme la probabilité d'avoir deux voxels d'intensité i et j , séparés par un vecteur (d, θ) correspondant à un déplacement $(\Delta x, \Delta y)$:

$$C_{\Delta x, \Delta y}(i, j) = \frac{1}{N \times M} \sum_{p=1}^{N-\Delta x} \sum_{q=1}^{M-\Delta y} \begin{cases} 1 & \text{si } I(p, q) = i \text{ et } I(p + \Delta x, q + \Delta y) = j \\ 0 & \text{sinon} \end{cases} \quad (2.1)$$

où $I(p, q)$ représente le voxel de coordonnées (p, q) dans l'image I de taille $N \times M$. Le plus souvent la distance d est choisie égale à 1. Le vecteur \vec{d} permet de former toutes les combinaisons possibles entre un voxel et ses 26 voisins en 3-dimensions. Sachant que seulement 13 directions sont indépendantes, il faudra donc calculer 13 matrices de co-occurrence. Un exemple du calcul de la matrice de co-occurrence est donné sur la figure 2.4.

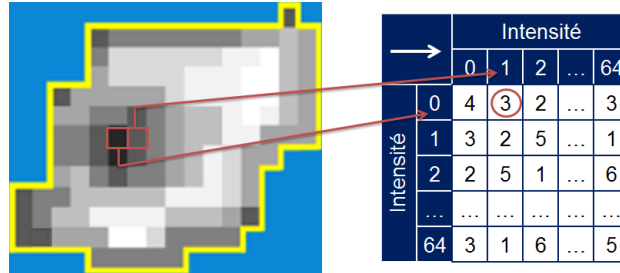


FIGURE 2.4 – Exemple de remplissage de la matrice de co-occurrence (avant normalisation par le nombre de voxels dans l'image) pour un vecteur de déplacement $\vec{d} = (1, 0)$. Le voxel de gauche entouré en rouge a une intensité de 0 et celui de droite une intensité de 1.

2.2.2.1.2 Calcul des index

À partir de la matrice de co-occurrence, il est possible d'extraire des index de texture appelés parfois caractéristiques d'Haralick :

- *Homogeneity* : mesure l'homogénéité de la distribution et vaut 1 pour une image constante.

$$\sum_i \sum_j \frac{C(i, j)}{1 + |i - j|} \quad (2.2)$$

- *Energy* (ou *Uniformity* ou *Angular Second Moment*) : mesure l'uniformité de la texture et augmente pour une distribution constante de niveaux de gris.

$$\sum_i \sum_j C(i, j)^2 \quad (2.3)$$

- *Correlation* : mesure la dépendance linéaire des niveaux de gris dans l'image.

$$\sum_i \sum_j \frac{(i - \mu_i) \times (j - \mu_j) \times C(i, j)}{\sigma_i \times \sigma_j} \quad (2.4)$$

- *Contrast* (ou *Inertia*) : mesure le contraste dans l'image et augmente avec celui-ci.

$$\sum_i \sum_j (i - j)^2 \times C(i, j) \quad (2.5)$$

- *Entropy* : mesure le désordre dans l'image et augmente lorsque la texture est aléatoire.

$$- \sum_i \sum_j C(i, j) \times \log(C(i, j)) \quad (2.6)$$

- *Dissimilarity* : mesure de la variation des niveaux de gris dans l'image.

$$\sum_i \sum_j |i - j| \times C(i, j) \quad (2.7)$$

- *Autocorrelation* : mesure les relations spatiales linéaires entre les tailles des différentes textures.

$$\sum_i \sum_j ij \times C(i, j) \quad (2.8)$$

- *Cluster Prominence* : mesure de l'assymétrie de la matrice de co-occurrence. Quand l'index est élevé, l'image n'est pas symétrique et quand il est bas le pic de la matrice de co-occurrence est autour de la valeur moyenne.

$$\sum_i \sum_j [i + j - \mu_i - \mu_j]^4 \times C(i, j) \quad (2.9)$$

- *Cluster Shade* : caractérise la manque de symétrie de la matrice.

$$\sum_i \sum_j [i + j - \mu_i - \mu_j]^3 \times C(i, j) \quad (2.10)$$

- *Cluster Tendency* : mesure le regroupement des voxels ayant des niveaux de gris similaires.

$$\sum_i \sum_j [i + j - \mu_i - \mu_j]^2 \times C(i, j) \quad (2.11)$$

- *Difference entropy* : mesure du désordre sur l'histogramme des différences entre les lignes et les colonnes de la matrice.

$$\sum_i C_{x-y}(i) \times \log(C_{x-y}(i)) \quad (2.12)$$

- *Informational measure of correlation - 1* : mesure la dépendance linéaire des niveaux de gris des voxels voisins.

$$\frac{H - HXY1}{\max\{HX, HY\}} \quad (2.13)$$

- *Informational measure of correlation - 2* : mesure la corrélation moyenne dans l'image.

$$\sqrt{1 - e^{-2(HXY2 - H)}} \quad (2.14)$$

- *Inverse Difference Moment Normalized* : mesure le degré de lissage de l'image.

$$\sum_i \sum_j \frac{C(i, j)}{1 + \left(\frac{|i-j|^2}{N^2}\right)} \quad (2.15)$$

- *Inverse Difference Normalized* : mesure la similarité des textures dans l'image.

$$\sum_i \sum_j \frac{C(i, j)}{1 + \left(\frac{|i-j|}{N}\right)} \quad (2.16)$$

- *Maximum Probability* : détecte le motif le plus fréquent.

$$\max\{C(i, j)\} \quad (2.17)$$

- *Sum average* : mesure la somme des voxels adjacents sur l'histogramme des sommes des lignes et des colonnes.

$$\sum_i i \times C_{x+y}(i) \quad (2.18)$$

- *Sum entropy* : mesure le désordre sur l'histogramme des sommes.

$$- \sum_i C_{x+y}(i) \times \log(C_{x+y}(i)) \quad (2.19)$$

- *Sum variance* : mesure la variance sur l'histogramme des sommes.

$$\sum_i (i - \text{Sum entropy})^2 \times C_{x+y}(i) \quad (2.20)$$

- *Variance* : mesure la variabilité des niveaux de gris.

$$\sum_i \sum_j (i - \mu)^2 \times C(i, j) \quad (2.21)$$

où :

- $C(i, j)$ correspond à l'élément (i, j) de la matrice de co-occurrence ;
- μ_i ou μ_j à la moyenne sur la $i^{\text{ème}}$ ligne ou la $j^{\text{ème}}$ colonne de la matrice ;
- μ à la moyenne des $C(i, j)$;
- σ_i ou σ_j à la variance sur la $i^{\text{ème}}$ ligne ou la $j^{\text{ème}}$ colonne ;
- $C_x(i) = \sum_j C(i, j)$ et $C_y(j) = \sum_i C(i, j)$;
- $C_{x+y}(k) = \sum_i \sum_j C(i, j)$ pour $i + j = k$ et $C_{x-y}(k) = \sum_i \sum_j C(i, j)$ pour $|i - j| = k$;
- $HXY1 = - \sum_i \sum_j C(i, j) \times \log(C_x(i) \times C_y(j))$;
- $HXY2 = - \sum_i \sum_j C_x(i) \times C_y(j) \times \log(C_x(i) \times C_y(j))$;
- H correspond à Entropy, HX à l'Entropy sur les lignes et HY à l'Entropy sur les colonnes.

Dans le cas d'une image en 3 dimensions, l'index de texture est calculé indépendamment sur les 13 matrices, puis moyenné.

2.2.2.2 Matrice de différence de niveaux de gris

2.2.2.2.1 Calcul de la matrice

La matrice de différence de niveaux de gris, de l'anglais *Neighborhood Gray-Level Different Matrix* (NGLDM), [Amadasun and King, 1989] rend compte des différences de niveaux de gris entre un voxel et ses 26 voisins. La matrice est de taille $G \times 2$ et se calcule directement en 3 dimensions. L'élément $(i, 1)$ de la matrice NGLDM (D) correspond à la probabilité d'apparition du niveau i et l'élément $(i, 2)$ correspond à :

$$D(i, 2) = \sum_p \sum_q \begin{cases} |\bar{M}(p, q) - i| & \text{si } I(p, q) = i \\ 0 & \text{sinon} \end{cases} \quad (2.22)$$

où $\bar{M}(p, q)$ représente la moyenne de l'intensité des 26 voisins du voxel de coordonnées (p, q) . La figure 2.5 illustre le calcul de la matrice NGLDM.

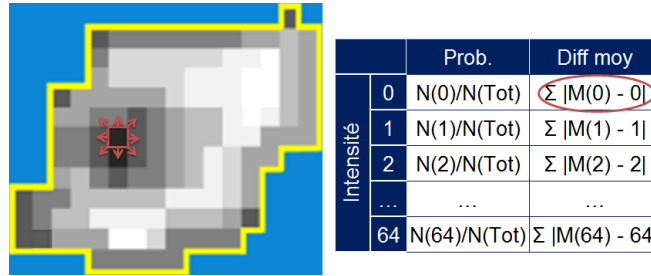


FIGURE 2.5 – Exemple du calcul de la matrice de différence de niveaux de gris. La moyenne des 26 voisins est calculée pour chaque voxel de laquelle on soustrait l'intensité du voxel central puis sommée sur l'ensemble des voxels de niveau i .

2.2.2.2.2 Calcul des index

À partir de la matrice NGLDM, on peut extraire cinq paramètres :

— *Coarseness* : mesure la rugosité d'une texture.

$$\frac{1}{\sum_i D(i, 1) \times D(i, 2)} \quad (2.23)$$

— *Contrast* : mesure le contraste sur l'image.

$$\left[\sum_i \sum_j D(i, 1) \times D(j, 1) \times (i - j)^2 \right] \times \frac{\sum_i D(i, 2)}{E \times G \times (G - 1)} \quad (2.24)$$

— *Busyness* : mesure les changements rapides d'intensité entre un voxel et ses voisins.

$$\frac{\sum_i D(i, 1) \times D(i, 2)}{\sum_i \sum_j (i \times D(i, 1) - j \times D(j, 1))} \text{ avec } D(i, 1) \neq 0, D(j, 1) \neq 0 \quad (2.25)$$

— *Complexity* : une texture est considérée comme complexe si le contenu d'information est élevé.

$$\sum_i \sum_j \frac{|i - j| \times (D(i, 1) \times D(j, 2) + D(j, 1) \times D(i, 2))}{E \times (D(i, 1) + D(j, 1))} \text{ avec } D(i, 1) \neq 0, D(j, 1) \neq 0 \quad (2.26)$$

- *Texture Strength* : mesure si les éléments qui composent l'image sont aisément définissables et clairement visibles.

$$\sum_i \sum_j \frac{(D(i, 1) + D(j, 1)) \times (i - j)^2}{\sum_i D(i, 2)} \text{ avec } D(i, 1) \neq 0, D(j, 1) \neq 0 \quad (2.27)$$

où E correspond au nombre de voxels dans le VOI et G au nombre de niveaux de gris.

2.2.2.3 Matrice des longueurs des séries homogènes

2.2.2.3.1 Calcul de la matrice

La matrice des longueurs des séries homogènes appelée aussi *Gray-Level Run Length Matrix* (GLRLM) [Xu et al., 2004] renseigne sur la longueur des séries ou des segments de voxels de même intensité dans l'image suivant un angle θ donné. La matrice GLRLM est une matrice de taille $G \times S$ avec G le nombre de niveaux de gris dans l'image après discrétisation et S la taille de la plus grande série homogène. L'élément (i, j) de la matrice GLRLM correspond au nombre de séries de j voxels ayant une intensité de i dont un exemple est donné sur la Figure 2.6. Comme pour la matrice de co-occurrence, pour rendre compte de la texture en 3 dimensions, la matrice GLRLM doit être calculée suivant 13 directions.

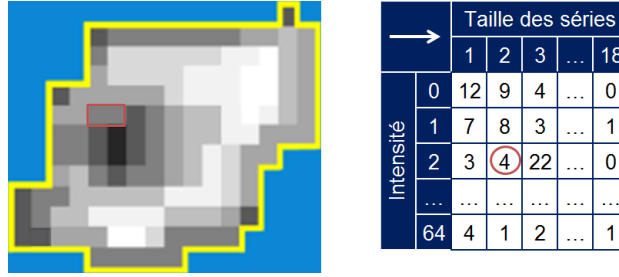


FIGURE 2.6 – Exemple de remplissage de la matrice des longueurs des séries homogènes suivant un angle θ égal à 0° . La série de voxels entourée en rouge correspond à une suite de 2 voxels d'intensité 2.

2.2.2.3.2 Calcul des index

À partir de la matrice GLRLM, 11 index de texture peuvent être calculés :

- *Short-Run Emphasis* (SRE) : mesure la distribution des courtes séries.

$$\frac{1}{H} \sum_i \sum_j \frac{R(i, j)}{j^2} \quad (2.28)$$

- *Long-Run Emphasis* (LRE) : mesure la distribution des longues séries (à l'inverse de SRE).

$$\frac{1}{H} \sum_i \sum_j R(i, j) \times j^2 \quad (2.29)$$

- *Low Gray-level Run Emphasis* (LGRE) : mesure la distribution des séries de bas niveaux de gris.

$$\frac{1}{H} \sum_i \sum_j \frac{R(i, j)}{i^2} \quad (2.30)$$

- *High Gray-level Run Emphasis* (HGRE) : mesure la distribution des séries de hauts niveaux de gris (à l'inverse de LGRE).

$$\frac{1}{H} \sum_i \sum_j R(i, j) \times i^2 \quad (2.31)$$

- *Short-Run Low Gray-level Emphasis* (SRLGE) : mesure la distribution des courtes séries de bas niveaux de gris.

$$\frac{1}{H} \sum_i \sum_j \frac{R(i, j)}{i^2 \times j^2} \quad (2.32)$$

- *Short-Run High Gray-level Emphasis* (SRHGE) : mesure la distribution des courtes séries de haut niveaux de gris.

$$\frac{1}{H} \sum_i \sum_j \frac{R(i, j) \times i^2}{j^2} \quad (2.33)$$

- *Long-Run Low Gray-level Emphasis* (LRLGE) : mesure la distribution des longues séries de bas niveaux de gris.

$$\frac{1}{H} \sum_i \sum_j \frac{R(i, j) \times j^2}{i^2} \quad (2.34)$$

- *Long-Run High Gray-level Emphasis* (LRHGE) : mesure la distribution des longues séries de haut niveaux de gris.

$$\frac{1}{H} \sum_i \sum_j R(i, j) \times i^2 \times j^2 \quad (2.35)$$

- *Gray-Level Non-Uniformity for run* (GLNUR) : mesure la non-uniformité des niveaux de gris.

$$\frac{1}{H} \sum_i \left(\sum_j R(i, j) \right)^2 \quad (2.36)$$

- *Run Length Non-Uniformity* (RLNU) : mesure la non-uniformité des longueurs des séries homogènes.

$$\frac{1}{H} \sum_j \left(\sum_i R(i, j) \right)^2 \quad (2.37)$$

- *Run Percentage* (RP) : mesure l'homogénéité de la distribution des séries homogènes.

$$\frac{H}{\sum_i \sum_j (j \times R(i, j))} \quad (2.38)$$

où $R(i, j)$ correspond à l'élément (i, j) de la matrice GLRLM et H au nombre de séries homogènes dans le VOI.

2.2.2.4 Matrice des longueurs des zones homogènes

2.2.2.4.1 Calcul de la matrice

La matrice des longueurs des zones homogènes ou *Gray-Level Zone Length Matrix* (GLZLM) [Thibault et al., 2009] correspond à une variante de la matrice précédente mais pour les zones homogènes. Elle ne nécessite pas de choix de la direction θ puisqu'elle est calculée directement en 3 dimensions. La matrice GLZLM est une matrice de taille $G \times T$ avec G le nombre de niveaux de gris dans l'image après discrétisation et T la taille de la plus grande zone homogène. L'élément (i, j) de la matrice GLZLM correspond au nombre de zones de j voxels ayant une intensité de i comme expliqué sur la figure 2.7.

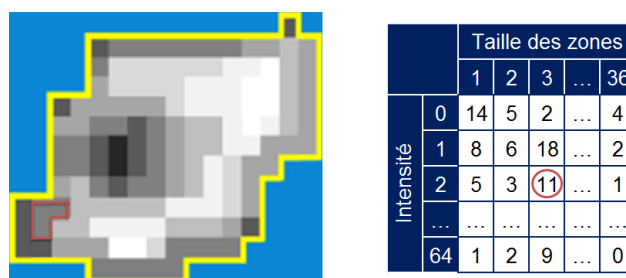


FIGURE 2.7 – Exemple d'une matrice des longueurs des zones homogènes. Les voxels entourés en rouge correspondent à une zone de 3 voxels d'intensité 2.

2.2.2.4.2 Calcul des index

Comme pour la matrice GLRLM, 11 index de texture peuvent être calculés à partir de GLZLM. Dans les équations précédentes, $R(i, j)$ devient $Z(i, j)$, l'élément (i, j) de la matrice GLZLM et H correspond au nombre de zones homogènes dans le VOI.

2.3 État de l'art en imagerie TEP

2.3.1 Introduction

Depuis 2009, 36 publications ont rapporté des analyses de texture à partir des images TEP (revue à partir de *Pubmed* le 28 avril 2015 avec les mots clés : "texture or textural and PET"), avec une augmentation majeure du nombre d'études à partir de 2013 (Figure 2.8). Parmi ces études, un tiers portent sur les cancers pulmonaires, les quatre autres types de cancers principalement étudiés étant les cancers de l'œsophage, du col de l'utérus, ORL ou mammaires (Figure 2.9).

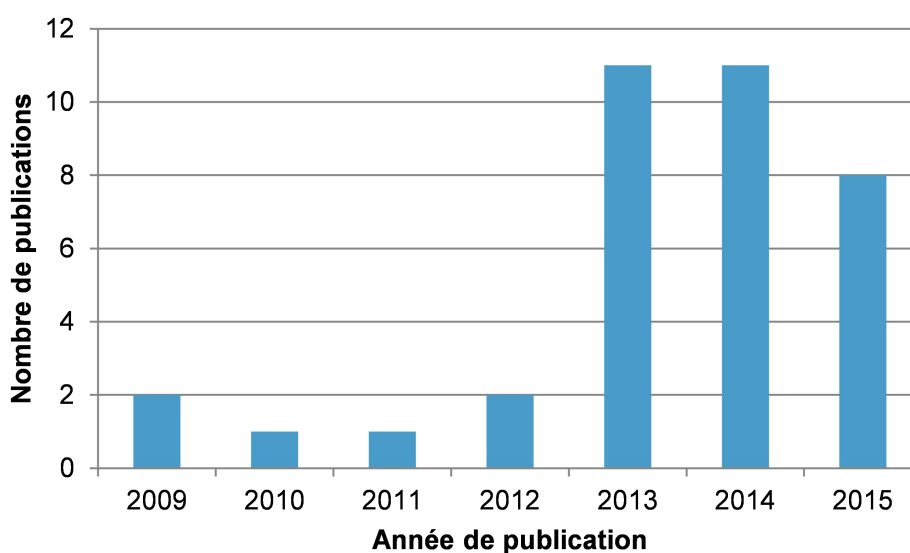


FIGURE 2.8 – Inventaire des publications portant sur l'évaluation de l'hétérogénéité tumorale à partir d'index de texture calculés sur des images TEP, suivant leur année de parution.

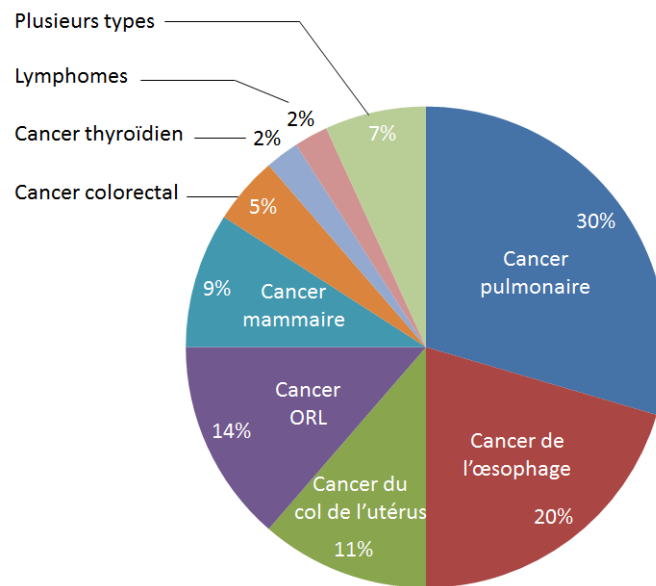


FIGURE 2.9 – Répartition des études portant sur l'analyse de texture à partir des images TEP suivant le type de cancer.

La grande majorité de ces publications (28 sur les 36 analysées) s'intéresse à l'impact clinique de l'analyse de texture pour :

- prédire la réponse à la thérapie, le risque de récurrence et/ou la survie ;
- segmenter un volume cible pour la radiothérapie ;
- développer des modules d'aide au diagnostic ;
- étudier la corrélation avec le stade clinique ;
- explorer la correspondance avec l'analyse visuelle ;
- comparer avec les caractéristiques histologiques ;
- étudier la corrélation avec les paramètres extraits des images TDM.

En parallèle, plusieurs articles de revue ou éditoriaux ([Kumar et al., 2012], [Asselin et al., 2012], [Davnall et al., 2012], [Visvikis et al., 2012], [Chicklore et al., 2013], [Houshmand et al., 2015]) ont porté sur l'analyse de texture en imagerie TEP démontrant l'intérêt croissant de la communauté pour cette technique mais aussi le besoin de comparer des études aux protocoles de calcul variés et avec des conclusions parfois contradictoires. Ces travaux font apparaître la nécessité de standardiser les méthodes de calcul des index de texture et d'étudier l'impact des différentes étapes sur les paramètres.

À partir des études répertoriées dans les Tableaux 2.1, 2.2 et 2.3, nous confrontons ces travaux suivant trois thèmes : l'influence des choix méthodologiques sur le calcul des index de texture, la robustesse des paramètres et enfin la comparaison des résultats obtenus dans les études cliniques.









Référence	Localisation	Nombre de tumeurs	Segmentation	F(D)	Matrice(s)
[El Naqa et al., 2009]	ORL, col de l'utérus	23	seuil fixe (40% SUVmax) et manuelle	?(?) 	CO
[Yu et al., 2009b]	ORL	10	but de l'étude	?(256) 	CO NGLDM
[Galavis et al., 2010]	plusieurs localisations	20	seuil fixe (40% SUVmax)	?(?) 	CO NGLDM GLRLM
[Tixier et al., 2011]	œsophage	41	FLAB	DR+1(16-128)	CO NGLDM GLRLM GLZLM
[Tixier et al., 2012]	œsophage	16	FLAB	DR+1(8-128)	CO GLZLM
[Vaidya et al., 2012]	poumon	27	manuelle	?(?) 	CO
[Tan et al., 2013a]	œsophage	20	seuil fixe (2.5 SUV)	?(64) 	CO
[Tan et al., 2013b]	œsophage	20	manuelle	?(64) 	CO
[Leijenaar et al., 2013]	poumon	45	seuil fixe (50% SUVmax) et manuelle	0,5 SUV près	CO GLRLM GLZLM
[Cheng et al., 2013]	ORL	70	seuil fixe (2,5 SUV)	?(4-64) 	CO NGLDM
[Hatt et al., 2013]	œsophage	50	seuil fixe (42% SUVmax), seuil adaptatif et FLAB	?(?) 	CO GLZLM
[Markel et al., 2013]	poumon	44	but de l'étude	0,05 SUV près	CO NGLDM

Tableau 2.1 – Première partie de l'état de l'art de l'analyse de texture à partir des images TEP. F : formule de discrétisation, D : taux de discrétisation, CO : matrice de co-occurrence, DR : discrétisation relative.







Référence	Localisation	Nombre de tumeurs	Segmentation	F(D)	Matrice(s)
[Bagci et al., 2013]	plusieurs localisations	30	but de l'étude	?(?) 	CO GLRLM
[Yang et al., 2013]	col de l'utérus	30	seuil fixe (40% SUVmax)	DR+1 (256)	GLRLM GLZLM
[Willaime et al., 2013]	sein	15	manuelle	DR (8-128)	CO NGLDM GLZLM
[Cook et al., 2013]	poumon	53	seuil fixe (45% SUVmax)	?(?) 	NGLDM
[Dong et al., 2013]	œsophage	40	seuil fixe (2,5 SUV)	?(?) 	CO
[Zhang et al., 2014]	œsophage	20	seuil fixe (2.5 SUV)	?(?) 	CO
[Oh et al., 2014]	ORL	70	seuil fixe (2,5 SUV)	?(?) 	NGLDM
[Tixier et al., 2014b]	poumon	102	FLAB	DR+1 (64)	CO GLZLM
[Soussan et al., 2014]	sein	54	seuil adaptatif	DR (64)	CO GLRLM
[Orlhac et al., 2014]	colorectal, sein, poumon	150	seuil fixe (40% SUVmax) et seuil adaptatif	DR(8-128) et DR+1(64)	CO NGLDM GLRLM GLZLM
[Brooks and Grigsby, 2014]	col de l'utérus	70 + simulations	seuil fixe (40% SUVmax)	?(255) 	CO
[Lartizien et al., 2014]	lymphomes	188	manuelle ajustée sur TDM	DR+1 (16)	CO NGLDM GLRLM GLZLM

Tableau 2.2 – Deuxième partie de l'état de l'art de l'analyse de texture à partir des images TEP. F : formule de discrétisation, D : taux de discrétisation, CO : matrice de co-occurrence, DR : discrétisation relative.









Référence	Localisation	Nombre de tumeurs	Segmentation	F(D)	Matrices(s)
[Xu et al., 2014]	plusieurs localisations	103	seuil adaptatif	?(?) 	CO NGLDM
[Tixier et al., 2014a]	colorectal	30	FLAB	DR+1 (64)	CO GLZLM
[van Gómez López et al., 2014]	poumon	38	seuil fixe (2,5 SUV)	?(?) 	CO
[Yip et al., 2014]	poumon	34	manuelle	DR (32)	CO NGLDM GLRLM
[Pyka et al., 2015]	poumon	45	seuil fixe (2,0 ou 2,5 SUV)	DR (64)	CO NGLDM
[Cook et al., 2015]	poumon	47	seuil fixe (40% SUVmax)	?(32) 	NGLDM
[Brooks and Grigsby, 2015]	col de l'utérus	85 + simulations	seuil fixe (40% SUVmax)	?(255) 	CO
[Lapa et al., 2015]	thyroïde	14	manuelle	?(?) 	CO GLZLM
[Hatt et al., 2015]	sein, col de l'utérus, ORL, poumon	555	FLAB	DR+1(4-256)	CO GLZLM
[Gao et al., 2015]	poumon	768	manuelle ajustée sur TDM	?(32) 	CO
[Dong et al., 2015]	œsophage, poumon	100	seuil fixe (2,5 SUV et 40% SUVmax) et manuelle	?(32) 	CO
[Cheng et al., 2015]	ORL	88	seuil fixe (2,5 SUV et 42% SUVmax) et seuil adaptatif	?(16-64) 	GLRLM GLZLM

Tableau 2.3 – Troisième partie de l'état de l'art de l'analyse de texture à partir des images TEP. F : formule de discrétisation, D : taux de discrétisation, CO : matrice de co-occurrence, DR : discrétisation relative.

2.3.2 Comparaison des choix méthodologiques pour le calcul des index

2.3.2.1 Différences dans la définition des index

La comparaison des publications portant sur l'analyse de texture des images TEP a permis de mettre en évidence des différences dans la définition des index. Par exemple, l'index Homogeneity n'obéit pas à la même formule dans [El Naqa et al., 2009] et dans [Tixier et al., 2011]. Dans la première référence, Homogeneity se calcule à partir de :

$$Homogeneity = \sum_i \sum_j \frac{C(i, j)}{1 + (i - j)^2} \quad (2.39)$$

alors que la seconde utilise :

$$Homogeneity = \sum_i \sum_j \frac{C(i, j)}{1 + |i - j|} \quad (2.40)$$

Des différences dans les formules mathématiques ont même été observées entre les publications d'un même groupe, comme Entropy pour lequel le signe moins disparaît entre [Tixier et al., 2011] et [Tixier et al., 2014a]. L'index ZLNU quant à lui répond à trois définitions totalement différentes entre [Tixier et al., 2011] :

$$ZLNU = \frac{1}{H} \sum_j \left[\sum_i Z(i, j) \right]^2 \quad (2.41)$$

[Tixier et al., 2012] :

$$ZLNU = \frac{1}{H} \sum_i \left[\sum_j \frac{Z(i, j)}{i^2} \right]^2 \quad (2.42)$$

et [Tixier et al., 2014a] :

$$ZLNU = \frac{1}{H} \sum_i \left[\sum_j Z(i, j) \right]^2 \quad (2.43)$$

D'autres publications ne définissent pas les index utilisés.

D'autre part, l'appellation des paramètres diffère entre les auteurs. Ainsi l'index que nous avons nommé ZLNU pour *Zone Length Non-Uniformity* peut prendre le nom de *Intensity Variability* dans [Tixier et al., 2011]. L'index Energy peut également être appelé *Uniformity* ou *Angular Second Moment*, tout comme Contrast qui peut répondre au nom de *Variance* ou *Inertia*. Quant à SZHGE, il peut être nommé *Szonehigl* [Willaime et al., 2013], *high-intensity short-zone emphasis* [Tixier et al., 2011] ou *high-intensity small-area emphasis* [Tixier et al., 2012].

Concernant le calcul des matrices, des variantes par rapport aux méthodes présentées au paragraphe 2.2.2.1.1 ont été proposées. Même si la majorité des études choisissent une distance d égale à 1 voxel pour le calcul de la matrice de co-occurrence, une équipe a proposé d'explorer en plus les résultats obtenus avec une distance de 3 voxels [Lartizien et al., 2014]. Les auteurs ont rapporté que les index calculés avec $d = 3$ conduisaient à de moins bonnes performances que ceux calculés avec $d = 1$ pour la distinction entre les zones hypermétaboliques cancéreuses et non-cancéreuses. Concernant cette même matrice de co-occurrence, récemment, des auteurs [Hatt et al., 2015] ont proposé de calculer directement la matrice en 3 dimensions et non plus par la moyenne de 13 directions indépendantes comme présenté au paragraphe 2.2.2.1.1.

2.3.2.2 Étape de discrétisation

Comme nous l'avons mentionné, une première étape, avant le calcul des matrices de texture, consiste à rééchantillonner les intensités des voxels présents dans la zone d'intérêt, afin de regrouper les voxels de même intensité et de limiter la taille des matrices à calculer. Pour ce faire, deux approches sont possibles :

- soit en effectuant une discrétisation relative (DR) qui consiste à rééchantillonner avec D valeurs discrètes l'intensité des voxels entre le minimum et le maximum d'intensité (SUV_{min} et SUV_{max}) dans le volume d'intérêt ;
- soit en arrondissant la valeur d'intensité du voxel suivant une largeur fixe comme par exemple à 0,5 unités SUV près [Leijenaar et al., 2013].

La majorité des études utilise la discrétisation relative mais la formule employée peut différer entre les auteurs d'un $+1$ au dénominateur :

$$R(x) = \text{arrondi}\left(D \times \frac{I(x) - SUV_{min}}{SUV_{max} - SUV_{min} + 1}\right) \quad (2.44)$$

$$R(x) = \text{arrondi}\left(D \times \frac{I(x) - SUV_{min}}{SUV_{max} - SUV_{min}}\right) \quad (2.45)$$

où $R(x)$ correspond à l'intensité du voxel x après discrétisation et $I(x)$ sa valeur initiale. Les conséquences sur les index de texture de l'utilisation de l'une ou l'autre formule seront détaillées au Chapitre 3 dans les contributions originales. Dans les Tableaux 2.1, 2.2 et 2.3, la première formule est choisie dans 7 cas (notée DR+1) et la seconde dans 4 cas (notée DR), notre étude de comparaison des deux méthodes n'étant pas comptabilisée.

De même, quand la discrétisation relative est choisie, la valeur de D varie selon les auteurs entre 4 et 256. L'impact de ce paramètre sera également plus amplement discuté au Chapitre 3. Toutefois parmi les 36 études présentées dans les Tableaux 2.1, 2.2 et 2.3, 22 ne rapportent pas la méthode utilisée et un tiers ne précisent pas la valeur de D choisie.

2.3.2.3 Volume minimum requis

Très rapidement, dans la littérature, la question du nombre minimum de voxels pour réaliser une analyse de texture fiable a été soulevée. Certains auteurs ont réalisé l'analyse de texture uniquement sur des volumes supérieurs à 3 cm^3 [Tixier et al., 2014b, Hatt et al., 2015], supérieurs à 4 cm^3 [Brooks and Grigsby, 2014] ou supérieurs à 5 cm^3 [Brooks and Grigsby, 2015], même si dans la majorité des cas, aucune exclusion volumique n'est rapportée dans les publications (19 publications sur 36). Pourtant, le volume tumoral semble influencer le calcul des index. [Brooks and Grigsby, 2014] ont montré que l'index Entropy reflète majoritairement le volume de la lésion et non l'hétérogénéité tumorale en dessous de 45 cm^3 .

2.3.3 Robustesse des index

Pour pouvoir utiliser l'analyse de texture et comparer les résultats entre les publications, il est nécessaire d'étudier la robustesse des index suivant les nombreux paramètres qui influent sur l'hétérogénéité de la fixation observée et la taille de la VOI.

2.3.3.1 Variations suivant les paramètres d'acquisition et de reconstruction

Dans leur article, [Galavis et al., 2010] étudient l'impact des paramètres d'acquisition et de reconstruction sur les index issus de la matrice de co-occurrence, de NGLDM et de GLRLM. Les auteurs ont utilisé les données de 20 patients pour lesquels ils ont considéré 10 types d'images en faisant varier le mode d'acquisition (2D ou 3D), la taille de la grille de voxels (128×128 ou 256×256), l'algorithme de reconstruction (OSEM ou *Iterative-View Point*), le nombre d'itérations (2 ou 4) et le post-filtrage (largeur du filtre de 3, 5 ou 6 mm). Un résumé des conclusions est présenté dans le Tableau 2.4. Les index Energy et LGRE sont peu variables suivant les paramètres d'acquisition et de reconstruction testés alors que Entropy, HGRE et GLNUR sont moyennement influencés. Tous les autres index testés présentent des variations supérieures à 30% suivant le type d'images. Néanmoins, cette étude ne détaille pas les causes de la grande variabilité de ces index (algorithme de reconstruction, grille de voxels ...).

2.3.3.2 Reproductibilité des index de texture

Dans un second temps, deux études [Tixier et al., 2012, Leijenaar et al., 2013] ont analysé la reproductibilité des index avec une analyse test-retest. Ce type d'étude consiste à acquérir, pour un même patient, deux examens TEP à quelques jours d'intervalle, afin de tester la stabilité des paramètres calculés à partir du calcul du coefficient de corrélation intra-classe ou ICC (abréviation de l'anglais *Intra-class Correlation Coefficient*). Une comparaison des deux études est présentée dans le Tableau 2.4. Tous les index issus de la matrice de co-occurrence apparaissent très reproductibles dans les deux études avec un $ICC \geq 80\%$ tout comme ZLNU, ZP et SUVmax ou SUVmean. Dans [Tixier et al., 2012], aucun index ne conduit à un ICC inférieur à 0,59 alors que dans [Leijenaar et al., 2013] l'index SZE conduit à un ICC égal à 0,35.

2.3.3.3 Sensibilité à la correction ou non de l'effet de volume partiel

Un autre facteur pouvant influencer l'analyse de texture provient de l'effet de volume partiel qui résulte à la fois du flou lié à la résolution spatiale et de l'échantillonnage grossier des images TEP. La correction ou non de cet effet de volume partiel peut donc modifier l'hétérogénéité de fixation du traceur visualisée sur les images. Jusqu'à présent, une seule étude a traité de cette question [Hatt et al., 2013]. Les auteurs montrent que les paramètres testés sont en majorité peu sensibles à la correction ou non de l'effet de volume partiel (cf. Tableau 2.4). Les index HGZE, GLNUz et le volume métabolique sont quant à eux moyennement influencés alors que ZLNU est fortement affecté avec une différence relative moyenne entre l'index issu des images avec et sans correction égale à $3,3 \pm 47,9\%$. Les capacités de prédiction de la réponse sont assez similaires entre les deux types d'images sauf pour ZLNU et HGZE qui bénéficient de la correction de l'effet de volume partiel (AUC respectivement égale à 0,72 et 0,65 sans correction contre 0,86 et 0,83 avec correction).

2.3.3.4 Impact du mouvement sur les index de texture

L'hétérogénéité tumorale peut également être influencée par la correction ou non des mouvements du patient. Dans leur étude, [Yip et al., 2014] se sont intéressés à la synchronisation ou non avec la respiration pour l'acquisition des images TEP de 26 patients ayant un cancer pulmonaire. Les auteurs ont montré que les index issus de la matrice NGLDM étaient significativement différents suivant l'acquisition des images avec ou sans synchronisation et que cette différence n'était pas influencée par le type histologique de la lésion (adénocarcinome vs. épidermoïde). Toutefois, on observe une corrélation modérée ($r \leq 0,70$) entre l'amplitude du mouvement rapportée au volume tumoral et la

	[Galavis et al., 2010]	[Tixier et al., 2012]	[Leijenaar et al., 2013]	[Hatt et al., 2013]
	Paramètres d'acquisition et de reconstruction	Test-retest	Test-retest	Correction de EVP
Matrice de co-occurrence				
Homogeneity	+++	+	+	+
Energy	+	+	+	∅
Correlation	+++	+	+	∅
Contrast	+++	+	+	∅
Entropy	++	+	+	+
Dissimilarity	+++	+	+	+
Matrice NGLDM				
Coarseness	+++	∅	∅	∅
Contrast	+++	∅	∅	∅
Busyness	+++	∅	∅	∅
Matrice GLRLM				
SRE	+++	∅	+	∅
LRE	+++	∅	+	∅
LGRE	+	∅	++	∅
HGRE	++	∅	+	∅
SRLGE	+++	∅	++	∅
SRHGE	+++	∅	+	∅
LRLGE	+++	∅	++	∅
LRHGE	+++	∅	+	∅
GLNUr	++	∅	++	∅
RLNU	+++	∅	+	∅
RP	+++	∅	+	∅
Matrice GLZLM				
SZE	∅	++	+++	∅
LZE	∅	+	++	∅
LGZE	∅	++	++	∅
HGZE	∅	+	+	++
SZLGE	∅	++	++	∅
SZHGE	∅	+	++	∅
LZLGE	∅	++	++	∅
LZHGE	∅	++	++	∅
GLNUz	∅	+	++	++
ZLNU	∅	+	+	+++
ZP	∅	+	+	+
Index conventionnels				
SUVmax	∅	+	+	∅
SUVmean	+++	+	+	+
SUVpeak	∅	∅	+	∅
TLG	∅	∅	∅	∅
VM	∅	∅	+	++

Tableau 2.4 – Comparaison des études de robustesse suivant les paramètres d'acquisition et de reconstruction, des tests-retests et la correction de l'effet de volume partiel (EVP). + : peu influencé, ++ : moyennement influencé, +++ : très influencé, ∅ : index non testé.

différence relative entre les index mesurés sur les images avec et sans synchronisation respiratoire. On peut donc se demander si la différence de texture mise en évidence entre les deux types d'images ne provient pas juste d'un niveau de bruit différent. De plus, l'impact de la synchronisation ou non n'a pas été testé sur la capacité des index à prédire la réponse au traitement et/ou la survie. Les index des trois autres matrices n'ont pas non plus été testés.

2.3.3.5 Robustesse vis-à-vis des méthodes de segmentation

Enfin, comme nous l'avons montré au Chapitre 1, il existe une grande variété de méthodes de segmentation des images TEP sans qu'aucune ne fasse consensus. Des auteurs ont donc étudié l'impact de la méthode de segmentation employée sur les valeurs des index de texture. En effet sur les études répertoriées dans les Tableaux 2.1, 2.2 et 2.3, 7 utilisent une segmentation manuelle, 16 un seuillage fixe soit de 2 ou 2,5 SUV, soit un pourcentage de SUVmax (40-50%), 4 publications utilisent un seuillage adaptatif et 6 une segmentation automatique nommée FLAB. [Leijenaar et al., 2013] ont étudié la stabilité des index suivant une segmentation manuelle réalisée par 5 experts sur 33 patients atteints de cancers pulmonaires. Ils ont observé une grande stabilité des index issus de la matrice de co-occurrence et de GLRLM avec un ICC supérieur à 0,80. Pour les index issus de GLZLM, 6 sont très robustes vis à vis des 5 segmentations manuelles, 3 le sont moyennement ($0,5 \leq ICC < 0,8$) et GL-NUz et SZLGE sont peu stables. Dans le cas des cancers ORL, [Cheng et al., 2015] se sont intéressés aux capacités prédictives de ZLNU suivant la méthode de segmentation : seuillage fixe (42% SUVmax) ou seuillage adaptatif. Sur 88 patients, ils ont montré que cet index avait une valeur pronostique pour les deux méthodes. Une troisième étude a porté sur la robustesse des index suivant 3 méthodes de segmentation : un seuillage fixe (42% SUVmax), un seuillage adaptatif et FLAB [Hatt et al., 2013]. Les auteurs ont étudié, pour 7 index de texture, l'évolution de l'AUC pour la prédiction de la réponse sur 50 patients atteints de cancer de l'œsophage. Ils ont observé que Homogeneity et Dissimilarity étaient statistiquement plus prédictifs avec une segmentation de type FLAB qu'avec une segmentation à seuil fixe d'après un test de Delong, mais aucun index ne conduit à une AUC significativement différente entre le seuillage adaptatif et FLAB. Les résultats complets des études de [Leijenaar et al., 2013] et [Hatt et al., 2013] sont résumés dans le Tableau 2.5. Une étude détaillée de la robustesse des index de texture suivant la méthode de segmentation sera présentée au Chapitre 3 dans les contributions originales.

2.3.4 Analyse des études cliniques

L'analyse de la littérature concernant l'intérêt des paramètres de texture en clinique montre tout d'abord que tous les auteurs ne calculent pas les mêmes index. Ainsi, la matrice de co-occurrence est calculée dans 31 cas sur 36, contre 14 pour la matrice NGLDM, 10 pour GLRLM et 13 pour GLZLM. Les tâches de classification sont également différentes entre études même si la principale concerne la prédiction de la réponse et/ou de la survie. Dans cette partie, nous allons donc confronter les études entre elles, relever les contradictions entre auteurs et présenter les facteurs qui limitent aujourd'hui l'utilisation de ces index en clinique.

2.3.4.1 Capacités prédictives des index de texture

2.3.4.1.1 Prédiction de la réponse

Dans le cas des cancers de l'œsophage, [Tan et al., 2013a] ont montré que l'index Contrast issu de la matrice de co-occurrence pouvait prédire la réponse à la thérapie sur un examen TEP post-

	[Hatt et al., 2013]	[Hatt et al., 2013]	[Leijenaar et al., 2013]
	FLAB vs. 42% SUVmax	FLAB vs. seuillage adaptatif	Segmentation manuelle par 5 experts
Matrice de co-occurrence			
Homogeneity	+++	++	+
Energy	∅	∅	+
Correlation	∅	∅	+
Contrast	∅	∅	+
Entropy	+	++	+
Dissimilarity	+++	+++	+
Matrice NGLDM			
Coarseness	∅	∅	∅
Contrast	∅	∅	∅
Busyness	∅	∅	∅
Matrice GLRLM			
SRE	∅	∅	+
LRE	∅	∅	+
LGRE	∅	∅	+
HGRE	∅	∅	+
SRLGE	∅	∅	+
SRHGE	∅	∅	+
LRLGE	∅	∅	+
LRHGE	∅	∅	+
GLNUr	∅	∅	+
RLNU	∅	∅	+
RP	∅	∅	+
Matrice GLZLM			
SZE	∅	∅	++
LZE	∅	∅	+
LGZE	∅	∅	++
HGZE	++	+++	+
SZLGE	∅	∅	+++
SZHGE	∅	∅	+
LZLGE	∅	∅	++
LZHGE	∅	∅	+
GLNUz	+++	+++	+++
ZLNU	+++	+++	+
ZP	++	++	+
Index conventionnels			
SUVmax	∅	∅	+
SUVmean	+++	+++	+
SUVpeak	∅	∅	+
TLG	∅	∅	∅
VM	+++	+++	+

Tableau 2.5 – Comparaison des études de robustesse suivant la méthode de segmentation. + : peu influencé, ++ : moyennement influencé, +++ : très influencé, ∅ : index non testé.

traitement chez 20 patients. Dans le même contexte, [Tixier et al., 2011] ont mis en évidence que Busyness et Contrast (NGLDM) ne pouvaient pas prédire la réponse, à la différence de tous les autres index de texture calculés chez 41 patients sur un examen pré-traitement.

Pour le cancer pulmonaire, Cook et al. ont d'abord montré sur 53 patients que Coarseness, Contrast (NGLDM) et Busyness étaient associés à la réponse aux traitements avec des AUC supérieures à 0,72 [Cook et al., 2013] alors que les auteurs, dans une seconde étude [Cook et al., 2015], ont rapporté que ces mêmes index n'étaient pas prédictifs (p -value > 0,13).

Concernant les cancers ORL, [El Naqa et al., 2009] ont montré sur seulement 9 patients que l'index Contrast (matrice de co-occurrence) pouvait séparer les patients répondeurs des patients non-répondeurs avec une AUC égale à 0,80. [Oh et al., 2014] ont quant à eux montré que Coarseness et Busyness étaient prédictifs de la réponse et non Contrast (NGLDM) chez 70 patients.

2.3.4.1.2 Prédiction de la survie

Pour des cancers pulmonaires non à petites cellules, Entropy apparaît prédictif de la survie globale sur 101 patients dans [Hatt et al., 2015] suivant une analyse univariée, et dans [Tixier et al., 2014b] qui utilisent la même cohorte à un patient près. À l'opposé, dans [Pyka et al., 2015], Entropy ainsi que les autres index testés ne sont pas prédictifs de la survie chez 45 patients. Dans [Cook et al., 2015], seul Contrast (NGLDM), mesuré sur les images TEP acquises 6 semaines après le début du traitement, est associé à la survie chez 40 patients.

Dans le cas des cancers ORL, [Cheng et al., 2015] montrent que ZLNU est prédictif de la survie sans progression chez 88 patients, les mêmes auteurs ayant précédemment démontré sur 70 patients que Energy était un facteur pronostique indépendant [Cheng et al., 2013]. Dans le même contexte, [Oh et al., 2014] ont montré que seuls Coarseness et Busyness étaient prédictifs de la survie globale sur 70 patients.

Pour les cancers thyroïdiens, une seule étude [Lapa et al., 2015] a montré qu'aucun des index calculés n'était prédictif de la survie globale sur 12 patients.

2.3.4.2 Corrélation avec les caractéristiques tumorales

Outre les capacités prédictives des index de texture, quelques études ont porté sur le lien entre ces paramètres et les caractéristiques des tumeurs. Ainsi [Dong et al., 2013] ont montré sur 40 patients atteints de cancer de l'œsophage, que les index Energy et Entropy étaient corrélés au stade tumoral T ainsi qu'au stade ganglionnaire N. [Willaime et al., 2013], dans la seule étude à ce jour avec le traceur 18F-FLT, ont mis en évidence que les tissus non pathologiques avaient un index Homogeneity plus élevé que les tissus mammaires tumoraux sur 11 patientes. L'évaluation du caractère bénin ou malin d'une lésion a également été étudiée par [Xu et al., 2014] qui ont montré que Energy, Entropy et Coarseness étaient capables de faire cette distinction sur 103 lésions. Enfin, [Lartizien et al., 2014] ont démontré que plusieurs index issus des images TEP, dont GLNUz, permettaient de différencier les zones hypermétaboliques cancéreuses et non-cancéreuses sur 188 régions segmentées.

2.3.4.3 Segmentation du volume cible en radiothérapie

Dans l'état de l'art établi dans les Tableaux 2.1, 2.2 et 2.3, trois articles s'intéressent à la délimitation du volume cible pour les plans de traitement en radiothérapie à l'aide de l'analyse de texture. Ainsi, [Yu et al., 2009b] montrent qu'un logiciel nommé COMPASS permet, à partir de l'analyse de

texture des images TEP et TDM, de séparer les voxels issus des tissus sains et cancéreux. Pour ce faire, les auteurs calculent la texture dans des fenêtres (par exemple $9 \times 9 \times 3$ voxels) et affectent au voxel central la valeur de l'index calculé. Le classement des voxels est ensuite guidé par un arbre de décision construit dans une étude précédente [Yu et al., 2009a]. Cette méthode automatique fournirait une délimitation précise dans le cas des cancers ORL et permettrait de réduire la variabilité inter-observateur observée quand cette segmentation est réalisée manuellement sur les images TDM et TEP au 18F-FDG. [Markel et al., 2013] utilisent également la texture mesurée dans des fenêtres de $7 \times 7 \times 3$ voxels à partir d'images TEP/TDM fusionnées pour améliorer la définition du volume cible pour les carcinomes pulmonaires. La méthode développée permet d'obtenir une sensibilité de 73,9% et une spécificité de 99,2% par rapport à un consensus de trois experts pris comme vérité terrain. Enfin, [Dong et al., 2015], pour des cancers du poumon et de l'œsophage, ont mis en évidence que l'hétérogénéité tumorale influençait la définition du volume cible. En effet, Entropy serait corrélé avec les différences de volume cible obtenus entre plusieurs méthodes utilisées en clinique. Les informations relatives à l'analyse de texture des lésions pourraient donc permettre d'améliorer la planification en radiothérapie.

2.3.4.4 Obstacles à une utilisation en clinique

Plusieurs raisons font qu'aujourd'hui les index de texture ne sont pas directement applicables en clinique. Outre le besoin de standardisation des protocoles de calcul pour améliorer la robustesse des index, davantage d'études cliniques sont nécessaires pour répondre aux questions soulevées dans la littérature et détaillées maintenant.

2.3.4.4.1 Comparaison avec les index classiques

Le premier point concerne la comparaison des performances des index de texture avec celles des index classiques qui sont utilisés en clinique. Même si la majorité des études compare les résultats de l'analyse de texture avec le SUVmax (30), seules 16 études reportent également les performances du volume métabolique. Au Chapitre 3, nous détaillerons l'impact que cela peut avoir sur les résultats.

2.3.4.4.2 Sens de variation des index

Les sens de variation des index entre les publications sont parfois contradictoires. Par exemple, Contrast (NGLDM) est plus élevé chez les patients répondeurs à la thérapie dans le cas de cancers pulmonaires [Cook et al., 2013] alors que l'index Contrast issu de la matrice de co-occurrence est plus élevé chez les patients non-répondeurs ayant un cancer de l'œsophage [Tan et al., 2013a]. Dans d'autres publications, les sens de variation des index sont contraires à l'intuition. [Tixier et al., 2014b] montrent que l'index Homogeneity augmente quand l'hétérogénéité tumorale augmente : les tumeurs classées hautement hétérogènes visuellement par deux experts ont donc un index Homogeneity plus élevé que les tumeurs classées homogènes. De même, Homogeneity et Entropy qui, par définition, varient en sens opposés, augmentent conjointement entre les patients répondeurs et non-répondeurs dans [Tixier et al., 2011].

2.3.4.4.3 Report des valeurs absolues

Jusqu'à présent, les études concernant l'analyse de texture des images TEP sont rétrospectives et la plupart des auteurs choisissent de ne pas présenter les valeurs absolues des index qui seront pourtant indispensables pour une utilisation prospective. D'autres choisissent de normaliser la valeur des index entre 0 et 1 sur leur cohorte, rendant toute comparaison entre études impossible

[Tixier et al., 2011, Tixier et al., 2014b]. L'identification de seuil par la réalisation de méta-analyses n'est donc actuellement pas réalisable.

2.3.4.4.4 Manque de fiabilité des conclusions

Outre la difficulté à tirer des conclusions d'études parfois contradictoires, une publication récente a également pointé le manque de fiabilité des analyses statistiques [Chalkidou et al., 2015]. Cette étude a repris les publications concernant l'analyse de texture en imagerie TEP de 2009 à 2013 afin de corriger les erreurs de type I, c'est-à-dire rejeter l'hypothèse nulle alors qu'elle est vraie, créant ainsi des faux positifs. Ce type d'erreur augmente quand on teste une multitude de variables sur une même population. Les auteurs ont montré qu'il y avait une probabilité moyenne de 76% que les études aient conclu à un résultat faussement positif. Ceci remet en cause les conclusions présentées précédemment lorsque qu'aucune correction des comparaisons multiples n'est effectuée et que les résultats ne sont pas validés sur un jeu de données indépendant.

Même si l'analyse de texture des images TEP pourrait apporter un bénéfice clinique en donnant des indications concernant la réponse aux traitements ou la survie du patient, trop de questions restent en suspens notamment vis à vis des discordances entre études et de la robustesse de ces paramètres. Dans les trois chapitres suivants, nous présenterons d'abord une étude méthodologique des index de texture, puis nous réaliserons une évaluation clinique sur trois cohortes de patients, avant d'étudier le lien entre l'hétérogénéité tumorale mesurée à différentes échelles.

Chapitre 3

Étude méthodologique des paramètres de texture

L'analyse de texture des images TEP connaît un véritable engouement depuis quelques années. Néanmoins, après une étude approfondie de la littérature décrite au Chapitre 2, les conclusions entre les auteurs sont parfois contradictoires. Avant d'utiliser les paramètres de texture dans un contexte clinique, il nous est donc apparu nécessaire d'étudier les différentes étapes de leur calcul et de comprendre l'impact qu'elles peuvent avoir sur les index. Dans ce chapitre, à partir de quatre cohortes de patients, d'acquisitions sur un fantôme et de simulations, nous avons étudié les relations entre les paramètres de texture et leur lien avec les index conventionnels, l'impact de la méthode de segmentation ainsi que le rôle de l'étape de discrétisation sur ces index. Enfin nous avons cherché à mieux comprendre leurs significations et variations en fonction d'hétérogénéités de fixation afin de donner des points de repères pour une utilisation future en clinique.

3.1 Matériels & méthodes

3.1.1 Patients

Pour réaliser cette étude, quatre cohortes ont été utilisées :

- 28 patients atteints de cancers colorectaux métastatiques (abrégiés CCM) ;
- 24 patients ayant un cancer pulmonaire non à petites cellules (CPNPC-1) ;
- 54 patientes ayant un cancer mammaire (CM) ;
- 48 patients ayant un cancer pulmonaire non à petites cellules (CPNPC-2).

La première cohorte a été constituée à l'Institut Jules Bordet de Bruxelles. Les patients ont été intégrés dans une étude clinique prospective qui avait pour but initial d'évaluer l'impact clinique de la réalisation d'un examen TEP au 18F-FDG avant et pendant le traitement par chimiothérapie [Hendlisz et al., 2012, Buvat et al., 2012]. L'étude a été approuvée par le comité d'éthique de l'Institut Jules Bordet et enregistrée sur clinicaltrials.gov avec le numéro NCT00741481. Le consentement écrit a également été obtenu pour tous les patients.

Les trois autres cohortes ont été traitées à l'Hôpital Avicenne de Bobigny. Cette étude a été approuvée par le comité local d'examen institutionnel (Ile-de-France-X) qui a levé l'obligation d'obtenir le consentement des patients. Les deux cohortes CPNPC-1 et CPNPC-2 ont été étudiées indépendamment. Dans la première, les lésions étaient localement avancées et ont nécessité un traitement néo-adjuvant complété ou non par une exérèse, alors que dans la seconde cohorte, les lésions étaient

localisées et les patients ont subi un traitement chirurgical d'emblée. Les caractéristiques de chaque cohorte sont résumées dans le Tableau 3.1.

Caractéristiques	CCM	CPNPC-1	CM	CPNPC-2
Sexe				
Homme	18	19	-	34
Femme	10	5	54	14
Age (années)				
Moyenne	63	62	55	64
Intervalle	27-83	48-79	31-83	48-85
Localisation				
Colon	2	-	-	-
Poumon	-	24	-	48
Sein	-	-	54	-
Foie	59	-	-	-
Péritoine	4	-	-	-
Autres	7	-	-	-
Stade clinique				
IA	-	-	-	12
IB	-	-	-	13
IIA	-	-	17	5
IIB	-	-	9	8
IIIA	-	16	2	9
IIIB	-	8	11	1
IIIC	-	-	1	-
IV	28	-	14	-
Type de lésion				
Primitive	2	24	54	48
Métastase	70	-	-	-

Tableau 3.1 – Caractéristiques des cohortes étudiées.

3.1.2 Acquisition des données TEP/TDM

Pour chaque patient, l'examen TEP au 18F-FDG avant traitement a été utilisé. Le taux de glucose sanguin était inférieur à 1,4g/L au moment de l'injection pour l'ensemble des patients.

Pour la cohorte issue de l'Institut Jules Bordet, les images TEP/TDM ont été acquises avec un imageur Discovery LS System (GE Healthcare), 69 ± 7 min (intervalle : [58–88 min]) après injection de 18F-FDG (4 MBq/kg) avec 4 min par pas de lit. Les images TEP ont été reconstruites avec le logiciel constructeur (GE Advance) en utilisant l'algorithme OSEM avec 2 itérations et 28 sous-ensembles avec un post-filtrage gaussien (largeur à mi-hauteur de 5,45 mm). Les images ont été corrigées de l'atténuation à partir de l'acquisition TDM et de la diffusion. Les images TDM ont été obtenues avec un scanner hélicoïdal multi-détecteurs à 4 coupes (Lightspeed, GE Healthcare). La tension était de 120 kV et le courant était compris entre 30 et 200 mA (déterminé automatiquement par l'algorithme Auto-mA de GE Healthcare). Les autres paramètres utilisés étaient : 0,5 s par rotation du TDM, un pitch de 1,5 et une vitesse de lit de 15 mm par rotation. La taille des voxels dans l'image reconstruite

est de 0,98 mm × 0,98 mm × 5,00 mm pour les images TDM et 3,91 mm × 3,91 mm × 4,25 mm pour les images TEP.

Pour les trois autres cohortes constituées à l'Hôpital Avicenne, les examens TEP/TDM ont été réalisés sur une machine Gemini temps de vol de Philips, 76±9 min (intervalle : [56–94 min]) après l'administration de 18F-FDG (3 MBq/kg) avec 1,45 min par pas de lit. Les images ont été reconstruites en utilisant l'algorithme BLOB-OS-TF en mode liste avec 2 itérations et 33 sous-ensembles. La correction de l'atténuation a été réalisée à partir des images TDM et aucun post-filtrage n'a été appliqué. Les images TDM ont été obtenues avec les réglages suivants : 120 kV, 100 mA, un pitch de 0,69. La taille des voxels sur les images reconstruites est de 1,17 mm × 1,17 mm × 1,5 mm sur les images TDM et de 4 mm × 4 mm × 4 mm sur les images TEP.

Les images TEP issues des deux centres ont été converties en unités SUV normalisées par la masse du patient (cf. paragraphe 1.2.1.2.1).

3.1.3 Segmentation du volume métabolique

Par défaut, la segmentation utilisée pour l'estimation du Volume Métabolique (VM) correspond à un seuillage adaptatif comme présenté au paragraphe 1.2.2.1. Après avoir effectué un contour manuel très grossier de la zone d'intérêt, nous appliquons un seuil égal à :

$$Seuil = \alpha I_{70} + I_{fond} \quad (3.1)$$

où le paramètre α (égal ici à 0,3) est déterminé à partir de trois acquisitions sur fantôme incluant des sphères de 0,98 à 3,12 cm de diamètre et avec un rapport entre l'activité des sphères et celle du fond variant de 2,96 à 10 [Maisonobe et al., 2013].

3.1.4 Analyse des images

Après segmentation, pour chaque lésion identifiée, deux types d'index ont été calculés : les index conventionnels, utilisés en clinique, et les index de texture.

3.1.4.1 Index conventionnels

Pour chaque VOI, cinq index conventionnels ont été extraits des images TEP avant traitement : SUVmax, SUVmean, SUVpeak, VM et TLG (cf. section 1.2).

3.1.4.2 Index de texture

Avant de calculer les matrices de texture, l'intensité des voxels contenus dans le VOI est d'abord discrétisée afin de réduire le bruit dans les images, de limiter la taille des matrices à calculer et de regrouper des voxels ayant des intensités proches. Par défaut, la méthode utilisée correspond à une discrétisation relative (DR) :

$$R(x) = \text{arrondi} \left(D \times \frac{I(x) - SUV_{min}}{SUV_{max} - SUV_{min}} \right) \quad (3.2)$$

où D est égal au nombre de valeurs discrètes fixé à 64 par défaut.

Quatre matrices de texture sont alors calculées : matrice de co-occurrence, NGLDM, GLRLM et GLZLM (cf. section 2.2). Trente-et-un index de texture sont extraits, correspondants à ceux sélectionnés dans [Tixier et al., 2011]. La liste de ces index est présentée dans le Tableau 3.2.

Matrice de co-occurrence
Homogeneity, Correlation, Contrast (CO), Dissimilarity, Energy, Entropy
Matrice NGLDM
Coarseness, Contrast (NGLDM), Busyness
Matrice GLRLM
SRE, LRE, LGRE, HGRE, SRLGE, SRHGE, LRLGE, LRHGE, GLNUr, RLNU, RP
Matrice GLZLM
SZE, LZE, LGZE, HGZE, SZLGE, SZHGE, LZLGE, LZHGE, GLNUz, ZLNU, ZP

Tableau 3.2 – Liste des index de texture calculés.

La segmentation, le calcul des index conventionnels et des index de texture ont été réalisés avec le logiciel Matlab, quant aux analyses statistiques présentées par la suite, elles ont été effectuées avec le logiciel R.

3.2 Étude de la corrélation entre index

Dans la première partie de cette étude méthodologique, nous avons étudié s'il existait une redondance entre les 31 index de texture calculés ou si chacun apportait une information différente.

3.2.1 Matériels & méthodes

À partir des cohortes CCM, CPNPC-1 et CM, nous avons calculé le coefficient de corrélation de Pearson (r) entre chaque paire d'index sur les trois cohortes indépendamment. Nous avons ensuite formé des groupes en réunissant les index ayant un coefficient r , en valeur absolue, supérieur ou égal à 0,80 sur au moins deux cohortes sur trois. Dans un second temps, nous avons calculé le coefficient de corrélation r entre les index de texture et SUVmax ou VM. Les caractéristiques du calcul des index de texture sont résumées ci-dessous.

Cohorte(s)	Segmentation	Taux (D)	Formule
CCM CPNPC-1 CM	Seuillage adaptatif	64	DR

3.2.2 Résultats

L'étude des index conventionnels et des index de texture nous a amené à calculer $(36 \times 35)/2 = 630$ coefficients de corrélation r pour chacune des trois cohortes séparément. En regroupant tous les patients, 281 coefficients r sont supérieurs ou égaux à 0,80, démontrant une très forte corrélation entre les index. Nous avons pu identifier 8 groupes d'index fortement corrélés entre eux deux à deux, qui sont présentés dans le Tableau 3.3. Seuls Busyness et LZLGE n'appartiennent à aucun groupe puisqu'ils ne présentent pas une forte corrélation avec les autres index, portant le nombre de groupe à 10.

Les résultats présentés dans le Tableau 3.4 montrent qu'aucun index de texture n'est fortement corrélé aux mesures SUV (SUVmax, SUVmean, SUVpeak) alors que 4 index de texture sont fortement corrélés au VM et au TLG. De plus, on observe une forte corrélation entre les index ayant la

Groupe	Index	CCM	CPNPC-1	CM
1	Homogeneity-Correlation-Contrast(CO) - Dissimilarity-Contrast(NGLDM)	0,94±0,04	0,90±0,07	0,91±0,06
2	Energy-Entropy-Coarseness	0,93±0,04	0,93±0,05	0,90±0,06
3	SRE-RP-SZE-ZP	0,98±0,01	0,89±0,08	0,97±0,02
4	LRE-LZE-LZHGE	0,87±0,09	0,48±0,44	0,89±0,05
5	LGRE-SRLGE-LRLGE-LGZE-SZLGE	0,98±0,02	0,64±0,32	0,97±0,02
6	HGRE-SRHGE-LRHGE-HGZE-SZHGE	0,96±0,03	0,91±0,09	0,96±0,04
7	GLNUr-RLNU-GLNUz-ZLNU-VM-TLG	0,97±0,03	0,94±0,07	0,91±0,07
8	SUVmax-SUVmean-SUVpeak	0,96±0,02	0,94±0,03	0,98±0,01
9	Busyness	-	-	-
10	LZLGE	-	-	-

Tableau 3.3 – Moyenne et écart-type de r sur chaque paire d'index présents dans les groupes de corrélation en fonction de la cohorte étudiée.

même définition mais issus de la matrice GLRLM et GLZLM comme par exemple SRE et SZE ou LGRE et LGZE. Le même constat peut être fait pour Contrast issu soit de la matrice de co-occurrence soit de NGLDM.

Finalement, nos résultats montrent une corrélation non négligeable entre le VM et les paramètres de texture sauf pour les index des groupes n°5, n°6, Busyness et LZLGE où $|r| < 0,64$ (cf. Tableau 3.4). À l'inverse, aucun index de texture n'est corrélé au SUVmax.

3.2.3 Discussion

En première approximation, nous avons considéré une corrélation linéaire entre les index et groupé ensemble ceux qui présentaient une forte corrélation suivant un seuil arbitraire fixé à 0,80. Nos résultats montrent que ces groupes de corrélation sont relativement indépendants du type de tumeur et des paramètres d'acquisition et de reconstruction. En effet, nous n'observons pas de tendance spécifique pour la cohorte CCM constituée à l'Institut Jules Bordet par rapport aux deux autres cohortes acquises à l'hôpital Avicenne. Les corrélations observées seraient donc plutôt inhérentes à la définition même des paramètres.

Comme suggéré dans [Tixier et al., 2011], nous retrouvons bien que les index issus des matrices GLRLM et GLZLM sont fortement corrélés (groupes n°3, 4, 5, 6, 7). De plus, aucun index de texture n'est fortement corrélé aux mesures SUV, ce qui est cohérent avec les résultats exposés dans [Dong et al., 2013, Hatt et al., 2013]. À l'inverse, la forte corrélation entre certains index de texture et le VM expliquent les résultats publiés précédemment pour les patients atteints de cancer de l'œsophage. GLNUz et ZLNU, pour lesquels nous avons montré une forte corrélation avec VM quelque soit la cohorte étudiée ($r > 0,93$), ont été identifiés comme prédictifs de la réponse [Tixier et al., 2011] alors que les résultats du VM n'étaient pas rapportés. Or, la même équipe avait déjà montré dans [Hatt et al., 2011] que le VM était un facteur prédictif de la réponse dans un groupe de patients similaire. On peut donc en déduire que le caractère prédictif de GLNUz et ZLNU mis en évidence dans [Tixier et al., 2011] n'était en fait dû qu'au caractère prédictif du volume métabolique, et non pas à une texture particulière de la fixation du 18F-FDG dans ces tumeurs œsophagiennes.

Groupe Index		SUVmax			VM		
		CCM	CPNPC-1	CM	CCM	CPNPC-1	CM
1	Homogeneity	0,16	0,32	-0,05	0,66	0,74	0,76
	Correlation	0,22	0,42	-0,12	0,67	0,73	0,64
	Contrast (CO)	-0,17	-0,35	0,07	-0,58	-0,72	-0,62
	Dissimilarity	-0,16	-0,36	0,08	-0,62	-0,75	-0,68
	Contrast (NGLDM)	-0,10	-0,31	0,14	-0,57	-0,71	-0,61
2	Energy	-0,21	-0,42	-0,18	-0,58	-0,54	-0,42
	Entropy	0,26	0,43	0,15	0,74	0,75	0,58
	Coarseness	-0,26	-0,43	-0,13	-0,74	-0,80	-0,70
3	SRE	-0,23	-0,28	-0,07	-0,76	-0,72	-0,82
	RP	-0,23	-0,23	-0,07	-0,76	-0,65	-0,83
	SZE	-0,21	-0,46	-0,02	-0,72	-0,81	-0,73
	ZP	-0,22	-0,43	-0,05	-0,74	-0,83	-0,80
4	LRE	0,24	0,16	0,07	0,77	0,56	0,83
	LZE	0,26	0,06	0,07	0,82	0,41	0,73
	LZHGE	0,31	0,55	0,19	0,86	0,86	0,76
5	LGRE	-0,05	-0,27	0,09	-0,48	-0,22	-0,42
	SRLGE	-0,05	-0,31	0,09	-0,48	-0,36	-0,43
	LRLGE	-0,04	-0,12	0,09	-0,47	0,12	-0,40
	LGZE	-0,06	-0,35	0,12	-0,48	-0,58	-0,44
	SZLGE	-0,09	-0,38	0,13	-0,47	-0,64	-0,45
6	HGRE	0,03	0,22	0,29	0,17	0,31	-0,13
	SRHGE	0,02	0,21	0,28	0,14	0,28	-0,15
	LRHGE	0,08	0,28	0,32	0,27	0,40	-0,03
	HGZE	0,02	0,17	0,23	0,28	0,27	-0,08
	SZHGE	-0,08	-0,04	0,13	-0,01	-0,07	-0,29
7	GLNUr	0,29	0,38	0,00	0,98	0,99	0,98
	RLNU	0,31	0,41	0,03	1,00	1,00	1,00
	GLNUz	0,30	0,38	-0,01	0,99	0,99	0,99
	ZLNU	0,30	0,36	0,05	0,99	0,97	0,93
	VM	0,31	0,41	0,02	1,00	1,00	1,00
	TLG	0,45	0,73	0,41	0,95	0,87	0,85
8	SUVmax	1,00	1,00	1,00	0,31	0,41	0,02
	SUVmean	0,94	0,96	0,97	0,14	0,32	-0,07
	SUVpeak	0,98	0,99	0,99	0,32	0,43	0,03
9	Busyness	-0,14	-0,06	0,25	-0,15	-0,13	-0,04
10	LZLGE	0,08	-0,06	0,02	-0,07	0,22	0,18

Tableau 3.4 – Coefficients de corrélation entre SUVmax ou VM et les index de texture en fonction de la cohorte étudiée. Les valeurs en gras sont supérieures ou égales à 0,80.

Quatre index appartenant au groupe n°1, Contrast issus de CO et de NGLDM, Homogeneity et Dissimilarity sont fortement corrélés ($|r| > 0,7$) dans notre étude, comme cela avait déjà été mis en évidence pour des images TEP au 18F-FLT pour des lésions mammaires [Willaime et al., 2013]. De même, dans le cas de cancers de l'œsophage, il a été rapporté que Homogeneity et Dissimilarity étaient hautement anti-corrélés avec $r = -0,93$ [Hatt et al., 2013].

Dans [El Naqa et al., 2009], Energy et Entropy (groupe de corrélation n°2) conduisent à des performances similaires pour prédire la réponse thérapeutique avec des AUC respectivement égales à 0,50 et 0,53 pour des cancers ORL et 0,72 et 0,65 pour des cancers du col de l'utérus. Ces résultats s'expliquent par une forte anti-corrélation entre ces deux index ($r < -0.92$). Dans la même publication, Homogeneity et Contrast (CO), issus eux aussi du même groupe de corrélation (n°1), présentent des performances de classification similaires dans les cancers ORL (AUC=0,80 et 0,82) et dans les cancers utérins (AUC=0,60 et 0,62).

Dans [Vaidya et al., 2012], les auteurs montrent que le coefficient de corrélation de Spearman (r_s) entre Energy et la récurrence tumorale loco-régionale ($|r_s| = 0,06$) est proche de celui obtenu entre Entropy et ce même critère de récurrence ($|r_s| = 0,02$). À partir de notre étude, nous avons donc montré que ce résultat s'explique par la très forte corrélation entre ces deux paramètres, qui appartiennent au groupe n°2. Le même constat peut être fait avec Contrast et Homogeneity.

Des performances similaires de prédiction de la réponse après une chimiothérapie néoadjuvante ont été rapportées pour Contrast et Correlation (AUC=0,85 et 0,80) [Tan et al., 2013a], ce résultat provenant du fait que les deux paramètres appartiennent au même groupe de corrélation (n°1). Dans [Dong et al., 2013], Energy et Entropy sont tous les deux fortement prédictifs du stade tumoral et ganglionnaire puisque ces deux paramètres sont également fortement corrélés (groupe n°2).

Nos résultats suggèrent donc que, dans les études cliniques, il n'y a pas de valeur ajoutée à calculer plusieurs index provenant du même groupe de corrélation. Néanmoins, ces conclusions nécessitent d'être validées sur d'autres types de tumeurs et pour d'autres machines TEP ou d'autres protocoles d'acquisition et/ou de reconstruction.

3.3 Impact de la méthode de segmentation

Comme nous l'avons vu précédemment, une majorité d'index de texture sont fortement ou modérément corrélés au VM. Puisque dans la littérature, aucune méthode de segmentation ne fait consensus, nous avons étudié l'impact de la méthode de segmentation sur les index de texture en comparant deux techniques souvent utilisées dans les publications : un seuillage fixe et un seuillage adaptatif.

3.3.1 Matériels & méthodes

Pour réaliser cette étude, nous avons utilisé les cohortes CCM, CPNPC-1 et CM. Pour étudier l'influence de la méthode de segmentation sur l'analyse de texture, nous avons segmenté à la fois les lésions avec un seuillage adaptatif (nommé VOI_{adapt}) et avec un seuillage fixe correspondant à 40% de la fixation maximale dans la tumeur ($VOI_{40\%}$). Les caractéristiques du calcul des index de texture sont résumées dans le tableau ci-dessous.

Cohorte(s)	Segmentation	Taux (D)	Formule
CCM CPNPC-1 CM	Seuillage adaptatif Seuillage fixe	64	DR

Dans un premier temps, nous avons comparé les VM obtenus à partir des deux méthodes de segmentation en réalisant un graphique de Bland-Altman sur les trois groupes de patients séparément. Dans un second temps, pour chaque type de tumeur, nous avons étudié la variation de l'index en

fonction de la méthode de segmentation, en évaluant la différence relative δ entre l'index mesuré à partir de $VOI_{40\%}$ et le même index issu de VOI_{adapt} :

$$\delta(\%) = \frac{Index(VOI_{40\%}) - Index(VOI_{adapt})}{(Index(VOI_{40\%}) + Index(VOI_{adapt}))/2} \times 100 \quad (3.3)$$

La moyenne et l'écart-type de δ (M_δ et sd_δ) sur toutes les tumeurs d'un même type ont ensuite été calculés pour caractériser la sensibilité de l'index à la méthode de segmentation utilisée.

3.3.2 Résultats

Les graphiques de Bland-Altman présentés sur la Figure 3.1 montrent qu'il existe une différence substantielle entre les VM issus des deux méthodes de segmentation. On note également que cette différence augmente quand le VM de la lésion augmente.

La moyenne et l'écart-type de δ calculés pour chaque index et chaque type de tumeur sont présentés dans le Tableau 3.5.

Une majorité de paramètres présente une bonne stabilité entre les deux méthodes de segmentation avec un écart moyen inférieur à 10% pour 21 index sur 31, comme SUVmax ou SUVpeak. Corrélation, SZHGE et LZLGE. SUVmean présente un écart moyen entre 10 et 20% alors que Energy, Coarseness, les index du groupe n°7 ainsi que Busyness varient de plus de 20% entre les deux méthodes de segmentation.

Aucun index de texture ne conduit à un sd_δ inférieur à celui de SUVmax ou SUVpeak, montrant que tous les paramètres de texture sont plus sensibles à la méthode de segmentation que SUVmax et SUVpeak. Seuls les sd_δ de 5 index de texture (Entropy, SRE, LRE, RP et SZE) sont inférieurs au sd_δ de SUVmean, pour les trois cohortes de patients. Douze index ont un sd_δ compris entre celui de SUVmean et du VM, alors que 14 paramètres conduisent à un sd_δ supérieur à celui de VM, montrant que ces index sont fortement dépendants de la méthode de segmentation.

3.3.3 Discussion

Aucune méthode de segmentation ne fait consensus dans la littérature. Pour une même tumeur, il est donc essentiel d'étudier la sensibilité des index de texture suivant le VOI utilisé pour les calculer. Pour ce faire, nous avons donc comparé deux méthodes fréquemment utilisées en clinique pour lesquelles nous obtenons des volumes tumoraux concordants (sans aberrations de segmentation, cf. Figure 3.1) mais avec des différences substantielles.

Cinq index de texture (Entropy, SRE, RP, SZE et LRE) diffèrent en moyenne très peu entre les deux méthodes et sont très peu variables tout comme SUVmax ou SUVpeak. Des index comme Contrast (CO ou NGDLM), LZE, LZHGE ou ceux du groupe n°5 sont eux peu influencés en moyenne mais les différences observées sont très variables entre les tumeurs ($sd_\delta(VM) < sd_\delta(index)$). D'autres paramètres comme Energy, Coarseness, Busyness ainsi que ceux du groupe n°7 diffèrent en moyenne de plus de 20% entre les deux segmentations et les différences observées sont également très variables d'une lésion à l'autre.

Ces résultats sont cohérents avec ceux présentés dans la littérature [Hatt et al., 2013] où les auteurs rapportaient, par exemple, que Entropy était peu ou moyennement influencé par la méthode de segmentation par une comparaison de trois méthodes (seuillage fixe à 42% de SUVmax, seuillage adapta-

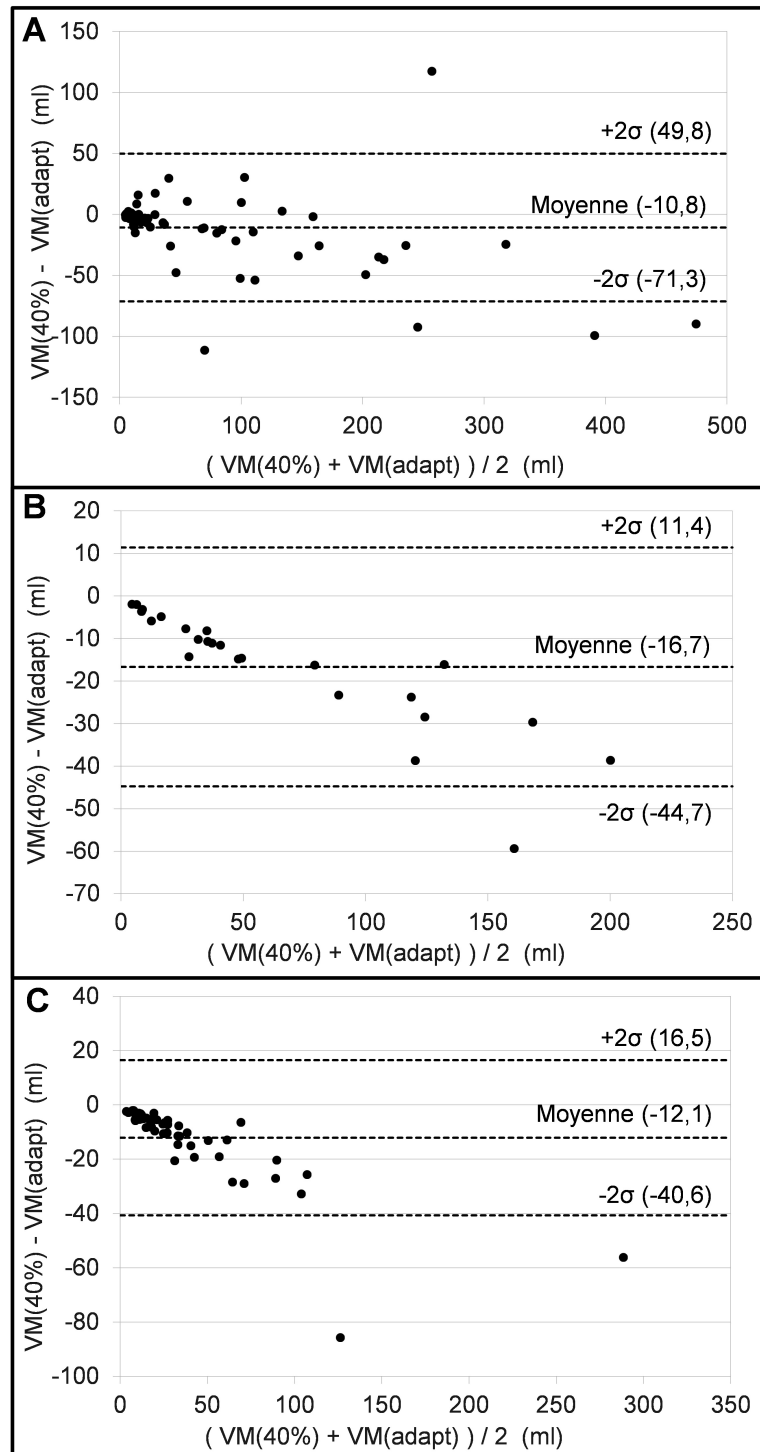


FIGURE 3.1 – Graphiques de Bland-Altman pour la comparaison des VM obtenus avec deux méthodes de segmentation ($VM_{40\%}$ et VM_{adapt}) pour la cohorte CCM (A), CPNPC-1 (B) et CM (C).

Groupe	Index	CCM	CPNPC-1	CM	Ecart moyen	Variabilité
1	Homogeneity	-3,2 ± 14,4	-0,0 ± 7,5	-5,3 ± 8,7	+	Modérée
	Correlation	-13,3 ± 27,4	-10,2 ± 18,1	-26,6 ± 36,4	++	Forte
	Contrast(CO)	5,6 ± 30,9	-3,5 ± 20,5	9,1 ± 19,7	+	Forte
	Dissimilarity	3,4 ± 18,2	-1,5 ± 10,7	5,3 ± 10,6	+	Modérée
	Contrast(NGLDM)	6,3 ± 36,2	-5,0 ± 24,6	8,1 ± 23,5	+	Forte
2	Energy	15,7 ± 31,1	29,4 ± 12,3	33,9 ± 16,5	+++	Forte
	Entropy	-3,5 ± 7,5	-5,6 ± 3,3	-7,2 ± 4,7	+	Faible
	Coarseness	13,1 ± 27,7	28,3 ± 11,1	26,0 ± 12,6	+++	Modérée
3	SRE	0,2 ± 0,6	0,1 ± 0,3	0,2 ± 0,3	+	Faible
	RP	0,2 ± 0,7	0,1 ± 0,4	0,3 ± 0,4	+	Faible
	SZE	1,7 ± 4,9	1,3 ± 3,5	2,4 ± 3,4	+	Faible
	ZP	3,1 ± 9,6	1,9 ± 5,3	4,1 ± 5,5	+	Modérée
4	LRE	-0,7 ± 2,2	-0,2 ± 1,3	-0,9 ± 1,1	+	Faible
	LZE	-9,5 ± 30,2	-6,4 ± 17,4	-13,5 ± 16,9	+	Forte
	LZHGE	-4,0 ± 21,8	3,5 ± 22,6	-6,5 ± 23,8	+	Forte
5	LGRE	-3,3 ± 51,9	6,3 ± 41,2	-9,7 ± 53,9	+	Forte
	SRLGE	-3,0 ± 51,8	6,6 ± 41,1	-9,2 ± 53,9	+	Forte
	LRLGE	-4,5 ± 52,5	4,8 ± 41,6	-11,5 ± 53,9	+	Forte
	LGZE	-0,9 ± 49,0	9,6 ± 39,0	-6,4 ± 53,3	+	Forte
	SZLGE	3,6 ± 53,3	10,4 ± 41,7	-3,0 ± 57,9	+	Forte
6	HGRE	5,1 ± 16,2	11,3 ± 10,7	7,8 ± 11,9	+	Modérée
	SRHGE	5,3 ± 16,3	11,5 ± 10,5	8,0 ± 11,7	+	Modérée
	LRHGE	4,5 ± 15,8	10,9 ± 11,3	6,7 ± 12,4	+	Modérée
	HGZE	4,8 ± 14,0	11,8 ± 10,1	8,4 ± 11,5	+	Modérée
	SZHGE	6,3 ± 15,3	13,0 ± 9,6	11,1 ± 11,2	++	Modérée
7	GLNUr	-17,4 ± 43,1	-24,3 ± 14,4	-32,1 ± 18,3	+++	Forte
	RLNU	-17,8 ± 37,4	-30,1 ± 9,1	-35,4 ± 13,0	+++	Modérée
	GLNUz	-14,5 ± 35,7	-22,3 ± 12,0	-28,2 ± 13,7	+++	Modérée
	ZLNU	-12,4 ± 26,5	-25,7 ± 9,7	-27,3 ± 11,8	+++	Modérée
	VM	-18,3 ± 38,8	-30,3 ± 9,6	-36,2 ± 13,6	+++	Modérée
	TLG	-12,2 ± 30,8	-19,4 ± 6,6	-22,9 ± 10,2	++	Modérée
8	SUVmax	0,3 ± 2,7	0 ± 0	0 ± 0	+	Faible
	SUVmean	6,7 ± 10,7	11,2 ± 4,4	13,7 ± 4,7	++	Faible
	SUVpeak	0 ± 0	0 ± 0	0 ± 0	+	Faible
9	Busyness	342,8 ± 4035,6	-113,7 ± 478,3	-527,9 ± 2403,8	+++	Forte
10	LZLGE	-15,2 ± 59,3	-6,9 ± 42,6	-27,4 ± 52,6	++	Forte

Tableau 3.5 – Moyennes (M_δ) et Ecart-types (sd_δ) de la différence relative entre un seuillage fixe et un seuillage adaptatif pour les trois cohortes séparément.

Trois catégories d'écart moyen ont été définies : + signifie $M_\delta(index) < 10\%$, ++ signifie $10\% \leq M_\delta(index) \leq 20\%$ et +++ signifie $M_\delta(index) > 20\%$.

Trois catégories de variabilité ont été définies : Faible signifie $sd_\delta(index) \leq sd_\delta(SUVmean)$, Modérée $sd_\delta(SUVmean) < sd_\delta(index) \leq sd_\delta(VM)$ et Forte $sd_\delta(VM) < sd_\delta(index)$.

tif et FLAB) alors que GLNUz et ZLNU (groupe n°7) sont très peu robustes. Toutefois, cette publication s'est limitée à l'examen de 7 index de texture sur les 31 initialement présentés [Tixier et al., 2011].

Dans leur étude, [Leijenaar et al., 2013] comparent cinq segmentations manuelles réalisées par des experts. Leurs résultats confirment la dépendance des index du groupe n°5 à la segmentation et la bonne robustesse d'autres paramètres comme Homogeneity, Entropy ou SRE.

Contrast et Busyness présentent une forte sensibilité à la méthode de segmentation. Or ces deux

index avaient déjà été rapportés comme fortement sensibles aux paramètres d'acquisition et de reconstruction [Galavis et al., 2010]. Cook et al., après avoir démontré que ces deux paramètres permettaient de prédire la réponse tumorale dans des lésions pulmonaires non à petites cellules [Cook et al., 2013] à partir d'une segmentation manuelle, ont mis en évidence récemment que Contrast et Busyness n'étaient pas capables de séparer les patients répondeurs ou non à la thérapie [Cook et al., 2015] à partir d'une segmentation par seuillage fixe (40% de SUVmax) dans le même type de tumeur. Les différences de performances entre ces deux dernières études peuvent s'expliquer par la faible robustesse de ces paramètres suivant la méthode de segmentation.

Comme pour l'étude des groupes de corrélation, la robustesse des index de texture suivant la méthode de segmentation ne dépend ni du type de tumeur, ni des protocoles d'acquisition et de reconstructions utilisés (cf. Tableau 3.5).

Pour la sélection des index à calculer, il sera donc préférable de choisir l'index le plus robuste à la méthode de segmentation dans chaque groupe, c'est-à-dire celui présentant le plus faible écart moyen et la plus faible variabilité. Pour le groupe n°1, on pourra choisir Homogeneity ou Dissimilarity qui ont un écart moyen faibles et sont modérément variables. Entropy et LRE apparaissent les plus robustes à la segmentation dans leur groupe respectif. Pour le groupe n°3, on pourra choisir parmi SRE, RP et SZE. Pour les groupes n°5, 6 et 7 tous les index présentent les mêmes caractéristiques vis-à-vis de la segmentation excepté SZHGE et GNLUr.

3.4 Impact de l'étape de discrétisation

L'étape de discrétisation est nécessaire avant le calcul des matrices de texture pour réduire le bruit dans les images, regrouper les voxels ayant des intensités similaires et limiter la taille des matrices à calculer. Cependant, cette étape diffère entre les auteurs tant pour le choix du taux de discrétisation que pour la formule utilisée. Nous allons donc ici en étudier l'impact sur le calcul des index de texture.

3.4.1 Taux de discrétisation

Comme nous l'avons présenté au paragraphe 2.3.2.2, le taux de discrétisation utilisé, quand il est spécifié par les auteurs, varie entre les publications. Nous avons donc cherché à en étudier les conséquences sur les index de texture.

3.4.1.1 Matériels & méthodes

Pour connaître l'impact du taux de discrétisation (D), nous avons calculé les index de texture avec une valeur D égale à 8, 16, 32, 64 et 128. Nous avons ensuite calculé les coefficients de corrélation de Pearson (r) entre les index calculés avec $D = 64$ et $D \neq 64$ (8, 16, 32 ou 128 valeurs discrètes) pour les trois cohortes de patients : CCM, CPNPC-1 et CM. Les caractéristiques du calcul des index de texture sont résumées dans le tableau ci-dessous.

Cohorte(s)	Segmentation	Taux (D)	Formule
CCM CPNPC-1 CM	Seuillage adaptatif	8-16-32-64-128	DR

3.4.1.2 Résultats

Les coefficients de corrélation de Pearson (r) entre les index calculés avec $D = 64$ et avec $D \neq 64$ sont présentés dans le Tableau 3.6 (moyenne sur les trois cohortes). Neuf index de texture (Homogeneity, Correlation, Contrast (CO et NGLDM), Dissimilarity, Coarseness, HGRE, GLNUR et RLNU) conduisent à un coefficient de corrélation supérieur à 0,95 pour toutes les comparaisons (8 vs. 64, 16 vs. 64, 32 vs. 64 et 128 vs. 64). Ce résultat suggère donc que ces paramètres sont très robustes vis-à-vis du taux de discrétisation utilisé. À l'inverse, Busyness et LZLGE sont fortement dépendants du taux de discrétisation avec r compris entre -0,01 et 0,84. Finalement, excepté pour ces deux index, tous les paramètres de texture conduisent à un coefficient de corrélation supérieur à 0,80 quand le taux de discrétisation est au moins égal à 32 valeurs discrètes, c'est-à-dire pour la comparaison 32 vs. 64 ou 64 vs. 128.

3.4.1.3 Discussion

L'étape de discrétisation est un préalable indispensable avant le calcul des index de texture et nous avons montré ici que le choix du taux de discrétisation affecte ces paramètres surtout en dessous de 32 valeurs discrètes pour des index comme Entropy ou HGZE. Tout comme [Tixier et al., 2012], nous retrouvons par exemple que LZE est influencé par la valeur de D .

Dans la littérature, la valeur de D varie entre 4 et 256 niveaux. Choisir une valeur de D trop faible revient à regrouper des voxels ayant des intensités éloignées, par exemple pour une tumeur variant entre 0 et 10 unités SUV et un $D=4$, il faut discrétiser les intensités à 2,5 unités près gommant une partie de la texture tumorale. À l'inverse, un taux trop élevé, comme $D=256$, sépare des voxels ayant

Groupe	Index	Coefficient de corrélation de Pearson (Moyenne±EC)			
		8 vs 64	16 vs 64	32 vs 64	128 vs 64
1	Homogeneity	0,98 ± 0,02	0,99 ± 0,01	1,00 ± 0,01	1,00 ± 0,01
	Correlation	1,00 ± 0,00	1,00 ± 0,00	1,00 ± 0,00	1,00 ± 0,00
	Contrast(CO)	1,00 ± 0,00	1,00 ± 0,00	1,00 ± 0,00	1,00 ± 0,00
	Dissimilarity	1,00 ± 0,00	1,00 ± 0,00	1,00 ± 0,00	1,00 ± 0,00
	Contrast(NGLDM)	0,98 ± 0,01	0,98 ± 0,01	1,00 ± 0,01	0,99 ± 0,00
2	Energy	-0,22 ± 0,19	0,54 ± 0,18	0,97 ± 0,02	1,00 ± 0,00
	Entropy	-0,29 ± 0,18	0,62 ± 0,13	0,96 ± 0,02	0,98 ± 0,01
	Coarseness	0,97 ± 0,01	0,98 ± 0,01	0,99 ± 0,00	0,99 ± 0,01
3	SRE	0,93 ± 0,08	0,95 ± 0,05	0,98 ± 0,02	0,97 ± 0,02
	RP	0,92 ± 0,09	0,95 ± 0,05	0,98 ± 0,01	0,97 ± 0,02
	SZE	-0,03 ± 0,11	0,59 ± 0,12	0,92 ± 0,02	0,93 ± 0,01
	ZP	0,81 ± 0,06	0,91 ± 0,04	0,97 ± 0,01	0,96 ± 0,01
4	LRE	0,93 ± 0,08	0,96 ± 0,03	0,98 ± 0,01	0,97 ± 0,02
	LZE	0,71 ± 0,34	0,75 ± 0,32	0,93 ± 0,07	0,97 ± 0,02
	LZHGE	0,85 ± 0,05	0,88 ± 0,05	0,92 ± 0,03	0,85 ± 0,11
5	LGRE	0,86 ± 0,05	0,93 ± 0,04	0,97 ± 0,01	0,98 ± 0,01
	SRLGE	0,88 ± 0,03	0,94 ± 0,03	0,97 ± 0,01	0,98 ± 0,01
	LRLGE	0,63 ± 0,32	0,93 ± 0,06	0,98 ± 0,02	0,98 ± 0,02
	LGZE	0,58 ± 0,09	0,91 ± 0,02	0,96 ± 0,01	0,97 ± 0,01
	SZLGE	0,41 ± 0,08	0,85 ± 0,05	0,91 ± 0,04	0,96 ± 0,01
6	HGRE	0,98 ± 0,01	1,00 ± 0,00	1,00 ± 0,00	1,00 ± 0,00
	SRHGE	0,95 ± 0,03	0,99 ± 0,01	1,00 ± 0,00	1,00 ± 0,00
	LRHGE	0,62 ± 0,17	0,90 ± 0,03	0,99 ± 0,01	1,00 ± 0,00
	HGZE	0,17 ± 0,26	0,52 ± 0,16	0,94 ± 0,01	0,98 ± 0,01
	SZHGE	0,02 ± 0,13	0,38 ± 0,11	0,86 ± 0,03	0,95 ± 0,02
7	GLNUr	1,00 ± 0,01	1,00 ± 0,00	1,00 ± 0,00	1,00 ± 0,00
	RLNU	0,98 ± 0,02	1,00 ± 0,01	1,00 ± 0,00	1,00 ± 0,00
	GLNUz	0,71 ± 0,12	0,86 ± 0,08	0,98 ± 0,02	1,00 ± 0,01
	ZLNU	0,68 ± 0,16	0,79 ± 0,21	0,97 ± 0,03	1,00 ± 0,01
9	Busyness	0,04 ± 0,07	0,09 ± 0,10	-0,01 ± 0,04	0,07 ± 0,07
10	LZLGE	0,30 ± 0,26	0,41 ± 0,47	0,61 ± 0,34	0,84 ± 0,14

Tableau 3.6 – Moyennes et écart-types (EC) sur les trois cohortes du coefficient de corrélation de Pearson entre l'index calculé avec $D = 64$ et le même index calculé avec $D \neq 64$. En gras les moyennes supérieures ou égales à 0,80.

une différence de fixation de 0,04 unités SUV dans notre exemple, bien en dessous de la précision avec laquelle les SUV sont estimés. Il convient donc de trouver un compromis, et comme dans la majorité des publications décrivant cette étape, nous avons donc choisi de fixer par défaut D égal à 64 valeurs discrètes dans la suite.

3.4.2 Méthode de discrétisation relative

Dans la littérature, l'étape de discrétisation, quand elle est spécifiée, correspond le plus souvent à une discrétisation relative. Toutefois, deux formules peuvent-être utilisées suivant les auteurs, nous avons donc cherché à en étudier l'impact sur le calcul des index de texture.

3.4.2.1 Matériels & méthodes

À partir des trois cohortes précédemment utilisées (CCM, CPNPC-1 et CM), nous avons comparé les index de texture calculés à partir de deux méthodes de discrétisation relative (cf. équation 3.2). Les résultats présentés jusqu'à présent dans ce manuscrit ont été obtenus avec la formule de discrétisation présentée dans l'équation 3.2. Or, dans la littérature, certains auteurs utilisent la formule de discrétisation relative proposée initialement par [Tixier et al., 2011] :

$$R(x) = \text{arrondi}\left(D \times \frac{I(x) - SUV_{min}}{SUV_{max} - SUV_{min} + 1}\right) \quad (3.4)$$

Nous avons donc calculé les index de texture à la fois avec la discrétisation 3.2 (nommée DR) et la discrétisation 3.4 (appelée DR+1 dans la suite). Les caractéristiques du calcul des index de texture sont résumées dans le tableau ci-dessous. Pour chaque index, nous avons ainsi calculé le coefficient de corrélation de Pearson (r) pour l'index obtenu avec les deux méthodes de discrétisation. Dans un second temps, nous avons calculé r entre l'index calculé avec DR+1 et SUVmax.

Cohorte(s)	Segmentation	Taux (D)	Formule
CCM CPNPC-1 CM	Seuillage adaptatif	64	DR et DR+1

3.4.2.2 Résultats

La comparaison entre les deux méthodes de discrétisation relative est présentée dans le Tableau 3.7. L'analyse des résultats montrent que 19 index de texture sur 31 sont très peu influencés par la méthode de discrétisation relative utilisée avec $r_{DR-DR+1} \geq 0,95$ sur les trois cohortes séparément. Dix paramètres sont moyennement influencés avec $0,80 \leq r_{DR-DR+1} < 0,95$ alors que deux index, Busyness et SZHGE sont plus fortement influencés $r_{DR-DR+1} = [-0,03-0,91]$.

Quand la formule DR+1 est utilisée, les coefficients de corrélation r des index du groupe n°6 avec SUVmax varient entre 0,27 et 0,67 (cf. Tableau 3.7). Dans le Tableau 3.4, nous avons observé que la corrélation entre les index du groupe n°6 calculés avec la méthode DR et SUVmax variait de -0,08 à 0,32, ce qui suggère donc que l'équation DR+1 introduit une corrélation avec SUVmax pour les index de ce groupe.

3.4.2.3 Discussion

L'utilisation de la formule 3.4 dans plusieurs publications [Tixier et al., 2011], [Yang et al., 2013] ou [Tixier et al., 2014a] entraînent une plus forte corrélation entre les index de texture du groupe n°6 et le SUVmax. Cette corrélation explique les résultats présentés précédemment par [Yang et al., 2013] qui montrent que 6 index de texture (HGRE, SRHGE, LRHGE, HGZE, SZHGE et LZHGE) calculés à partir de la méthode DR+1, ainsi que SUVmax et SUVmean sont corrélés avec les changements tumoraux observés dans le groupe des patients répondeurs.

Groupe	Index	Coefficient de corrélation de Pearson					
		CCM		CPNPC-1		CM	
		$r_{DR-DR+1}$	r_{SUVmax}	$r_{DR-DR+1}$	r_{SUVmax}	$r_{DR-DR+1}$	r_{SUVmax}
1	Homogeneity	0,99	0,04	1,00	0,24	0,95	-0,30
	Correlation	1,00	0,22	1,00	0,43	1,00	-0,12
	Contrast(CO)	0,98	-0,02	0,99	-0,26	0,95	0,34
	Dissimilarity	0,98	-0,03	0,99	-0,27	0,95	0,33
	Contrast(NGLDM)	0,99	-0,04	1,00	-0,27	0,98	0,28
2	Energy	1,00	-0,23	1,00	-0,42	1,00	-0,24
	Entropy	1,00	0,28	1,00	0,44	1,00	0,21
	Coarseness	1,00	-0,25	0,99	-0,44	1,00	-0,14
3	SRE	0,98	-0,13	0,98	-0,26	0,94	0,17
	RP	0,98	-0,13	0,99	-0,21	0,94	0,16
	SZE	0,96	-0,11	0,90	-0,32	0,92	0,24
	ZP	0,98	-0,12	0,96	-0,36	0,94	0,19
4	LRE	0,98	0,13	0,99	0,14	0,94	-0,16
	LZE	0,96	0,15	1,00	0,05	0,98	-0,02
	LZHGE	0,99	0,35	0,99	0,54	0,99	0,18
5	LGRE	0,99	-0,11	1,00	-0,29	0,98	-0,02
	SRLGE	0,99	-0,11	1,00	-0,33	0,99	-0,02
	LRLGE	0,99	-0,11	1,00	-0,12	0,98	-0,03
	LGZE	0,99	-0,13	1,00	-0,36	0,98	-0,01
	SZLGE	0,98	-0,14	1,00	-0,38	0,96	0,02
6	HGRE	0,92	0,37	0,97	0,42	0,88	0,65
	SRHGE	0,91	0,36	0,97	0,40	0,88	0,65
	LRHGE	0,93	0,39	0,98	0,45	0,89	0,66
	HGZE	0,90	0,38	0,95	0,41	0,81	0,67
	SZHGE	0,84	0,35	0,91	0,27	0,77	0,64
7	GLNUr	1,00	0,27	1,00	0,36	1,00	-0,05
	RLNU	1,00	0,31	1,00	0,41	1,00	0,03
	GLNUz	1,00	0,28	1,00	0,38	1,00	-0,04
	ZLNU	1,00	0,33	1,00	0,39	0,99	0,16
9	Busyness	-0,06	-0,07	-0,03	0,06	0,06	-0,01
10	LZLGE	0,90	-0,09	1,00	-0,06	0,92	-0,15

Tableau 3.7 – Coefficients de corrélation de Pearson (r) pour les index de texture calculés avec les deux formules de discrétisation relative et les index calculés avec la méthode DR+1 et SUVmax pour les trois cohortes séparément.

De plus, la méthode DR+1 ne correspond pas à un rééchantillonnage des intensités dans la tumeur avec D valeurs discrètes entre SUVmin et SUVmax. Par exemple, pour une tumeur dont la fixation varie entre 1,6 et 10,1 unités SUV (cf. Figure 3.2), pour un taux de discrétisation D égal à 64, la méthode DR discrétise bien les intensités entre 0 et 64 alors que la méthode DR+1 discrétise entre 0 et 58. De même, dans le cas d'un fantôme acquis sur la même machine TEP et dont la fixation varie entre 0,8 et 1,2 unités SUV, la méthode DR+1 discrétise entre 0 et 19 (cf. Figure 3.3).

Enfin, la corrélation avec SUVmax observée avec DR+1 provient de la discrétisation. En effet, avant discrétisation, deux tumeurs ayant la même gamme de variations de SUV (par exemple entre 2 et 6 unités SUV ou entre 6 et 10 unités SUV, soit une variation de 4 unités SUV) seront discrétisées entre 0 et 51 après DR+1. Or, en pratique, la gamme de variations des SUV dans une tumeur est très souvent corrélée au SUVmax de la dite tumeur.

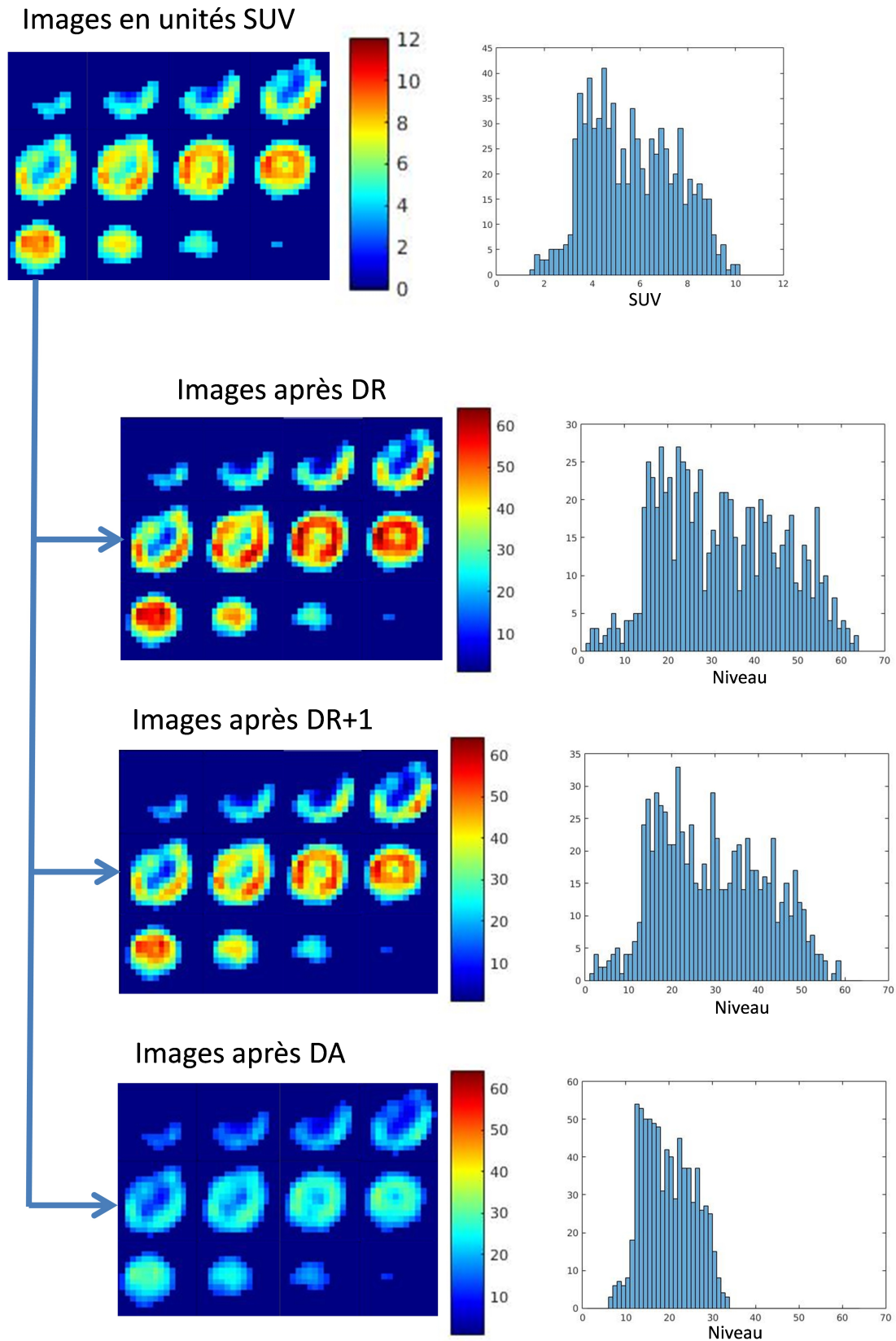


FIGURE 3.2 – Exemple des coupes d’une lésion vues en unités SUV, après DR, après DR+1 et après DA avec les histogrammes correspondants (pour $D=64$ valeurs discrètes).

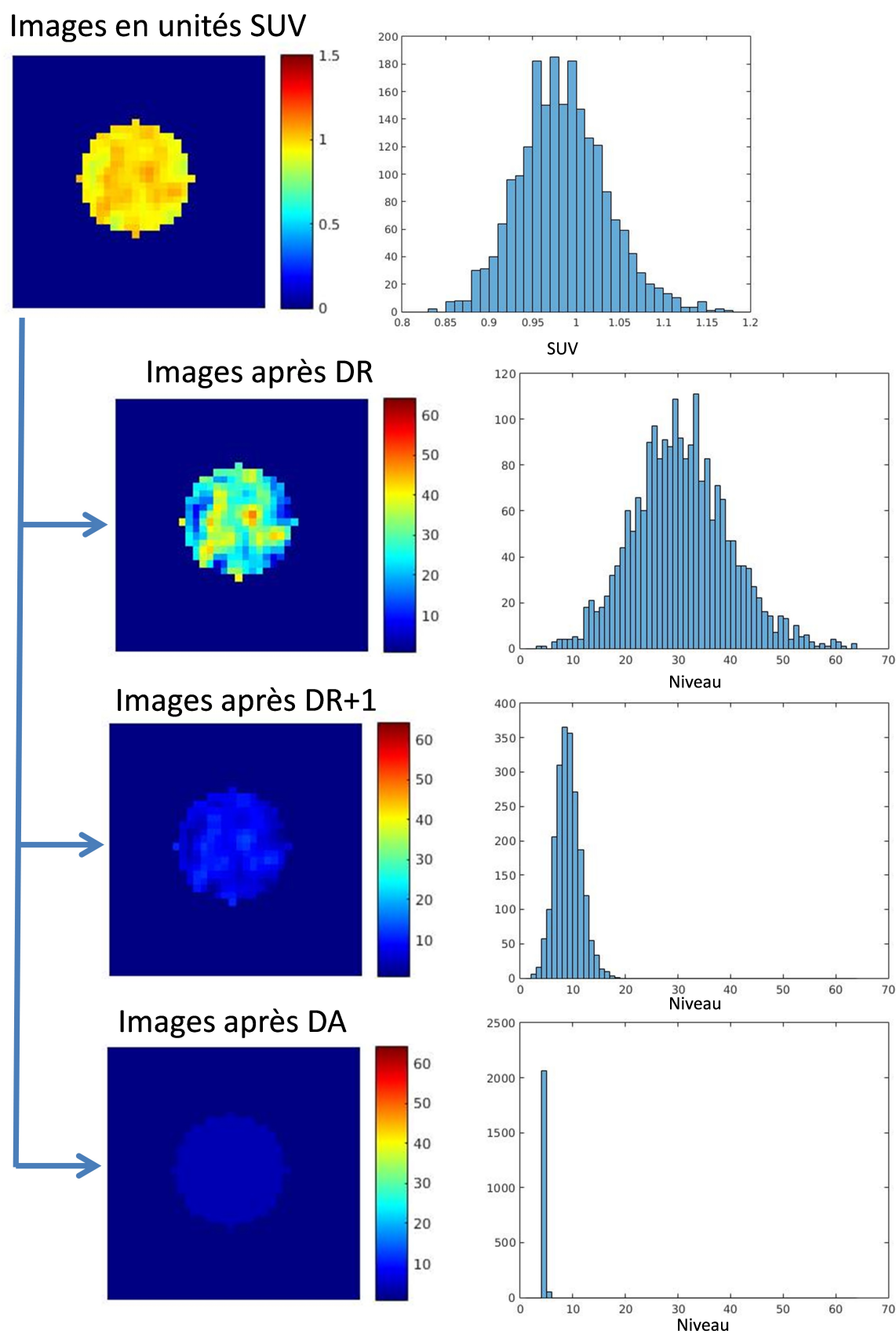


FIGURE 3.3 – Exemple de VOI tracé dans un fantôme vu en unités SUV, après DR, après DR+1 et après DA avec les histogrammes correspondants (pour $D=64$ valeurs discrètes).

3.4.3 Méthode de discrétisation absolue

En 2014, une étude présentée par [Brooks and Grigsby, 2014] a cherché à analyser le lien entre l'index de texture Entropy et le VM. À partir de la réalisation de simulations issues de tumeurs du col de l'utérus, les auteurs ont montré que Entropy représentait majoritairement le VM de lésion et non l'hétérogénéité de fixation du traceur en dessous de 45 cm³. Ce résultat a mis en lumière une limitation de l'analyse de texture en imagerie TEP. Toutefois, ces conclusions ont été obtenues sur un seul index. Nous avons donc cherché à vérifier ce résultat sur d'autres index de texture et avec une approche de discrétisation alternative.

3.4.3.1 Matériels & méthodes

Dans cette partie, nous avons utilisé les données provenant de la cohorte CPNPC-2 et des données acquises sur fantôme. Ce dernier est composé d'un cylindre de 20 cm de haut et de 16 cm de diamètre rempli d'eau dans laquelle ont été introduit 103,3 MBq de 18F-FDG. Pour l'acquisition des images, l'axe du cylindre est placé parallèlement à la direction axiale de l'imageur TEP/TDM, le même que celui utilisé pour la cohorte CPNPC-2 (Hopital Avicenne, Bobigny). Pour l'acquisition TEP, une seule position de lit est utilisée, d'une durée de 3 min. Les autres paramètres d'acquisition et de reconstruction sont les mêmes que ceux détaillés pour les acquisitions des données patients au paragraphe 3.1.2.

À partir des données issues de la cohorte CPNPC-2 et des données acquises sur un fantôme, nous avons étudié le lien entre les index de texture et les index conventionnels (SUVmax, VM) en fonction de la méthode de discrétisation utilisée.

Sur les données issues du fantôme, nous avons délimité automatiquement 5 sphères (nommées VOI-S dans la suite) et centrées sur 5 positions différentes comme présenté sur la Figure 3.4-A. Le diamètre des sphères varie entre 3 et 17 voxels (seuls les valeurs impaires sont utilisées afin de centrer la sphère sur un voxel).

Sur les données CPNPC-2, la lésion primitive est segmentée avec un seuillage adaptatif. Le VOI, nommé VOI-T dans la suite, est ensuite déplacé dans le foie du patient (VOI-F), que nous considérons comme une région de référence dans un tissu sain (cf. Figure 3.4-B).

Avant le calcul des matrices de texture, nous avons discrétisé les intensités des voxels dans VOI-S, VOI-T et VOI-F en utilisant deux méthodes de discrétisation :

- la discrétisation relative (DR) qui consiste à rééchantillonner la valeur des voxels entre SUVmin et SUVmax de chaque VOI (cf. équation 3.2),
- la discrétisation absolue (DA) qui consiste à discrétiser la valeur des voxels entre deux bornes fixes pour l'ensemble des VOI. Nous avons fixé ces bornes à 0 et 20 unités SUV, ce qui correspond à l'intervalle caractéristique de variations du SUV chez les patients :

$$R(x) = \text{arrondi}\left(D \times \frac{I(x)}{20}\right) \quad (3.5)$$

Les caractéristiques du calcul des index de texture sont résumées dans le tableau ci-dessous.

Cohorte(s)	Segmentation	Taux (D)	Formule
CPNPC-2 Fantôme	Seuillage adaptatif ø=3-17 voxels	64	DR et DA

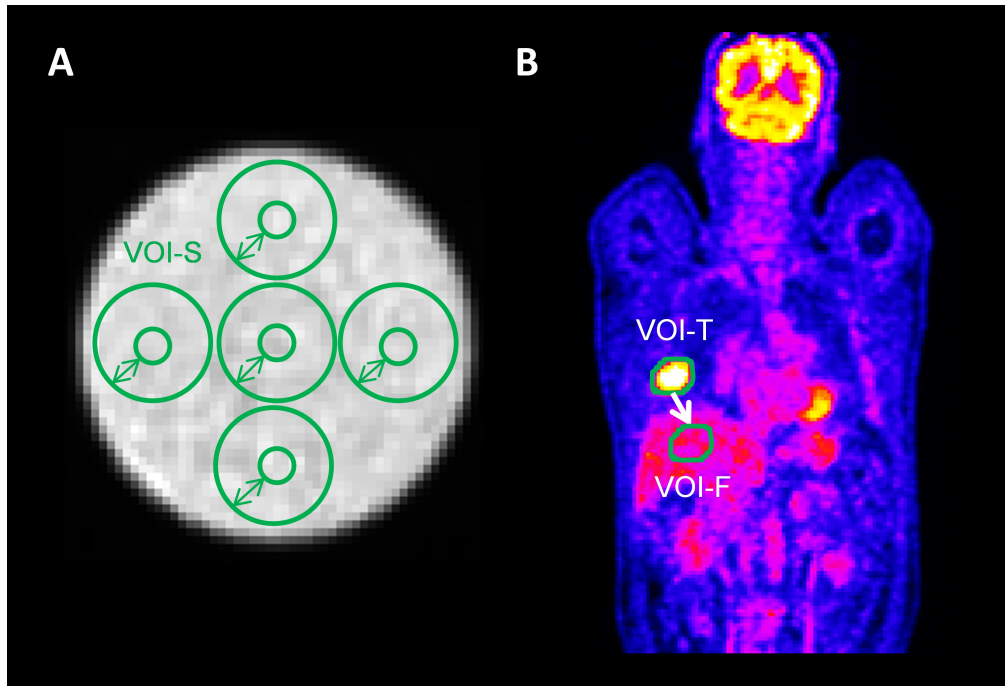


FIGURE 3.4 – Segmentation de sphères dans l'image TEP acquise sur un fantôme cylindrique, nommées VOI-S (A). Chez les patients, segmentation de la lésion tumorale primitive (VOI-T) avec un seuillage adaptatif puis report du volume dans le foie (VOI-F) (B).

Pour un index par groupe de corrélation, choisi parmi les plus robustes à la méthode de segmentation dans chaque groupe, nous avons tracé la valeur de l'index en fonction du VM pour VOI-S, VOI-T et VOI-F et en fonction de SUVmax pour les deux méthodes de discrétisation (DR et DA). Nous avons ensuite mesuré la corrélation entre les index et VM ou SUVmax à l'aide du coefficient de corrélation de Spearman (r_s) pour VOI-T et VOI-F.

3.4.3.2 Résultats

Après segmentation par un seuillage adaptatif, le volume tumoral moyen était de $19,5 \pm 26,1$ ml (intervalle : [2,8–119,0 ml]) dans la cohorte de patients CPNPC-2.

Relation entre les index de texture calculés avec la discrétisation relative et le VM ou SUVmax

L'évolution des index de texture en fonction du VM (cf. Figure 3.5) montre que la plupart des paramètres croissent ou décroissent au début de la courbe avant d'atteindre un plateau. Cette tendance est observée à la fois dans le fantôme, les tumeurs pulmonaires et le tissu sain. Les graphiques de Homogeneity, Entropy, SRE, LRE et LGZE (cf. Figure 3.5) montrent que ces paramètres sont fortement corrélés au VM ($|r_s|=[0,76-1]$, cf. Tableau 3.8) surtout pour de petits VOI. HGZE présente un autre comportement avec des larges fluctuations pour les petits VOI (cf. Figure 3.5). RLNU croît linéairement avec le VM sans jamais atteindre de plateau ($r_s=1$).

Les graphiques des index de texture en fonction de SUVmax présentés sur la Figure 3.6 montrent que les paramètres calculés à partir de la discrétisation relative sont indépendants de la fixation du traceur dans le VOI ($r_s=[0,02-0,36]$, cf. Tableau 3.8).

Index	Discrétisation Relative				Discrétisation Absolue			
	VOI-T		VOI-F		VOI-T		VOI-F	
	VM	SUVmax	VM	SUVmax	VM	SUVmax	VM	SUVmax
Homogeneity	0,87*	0,25	0,88*	0,14	-0,02	-0,90*	0,31*	-0,66*
Entropy	1,00*	0,36*	1,00*	0,03	0,75*	0,80*	0,22	0,72*
SRE	-0,80*	-0,36*	-0,88*	-0,08	-0,15	0,79*	-0,67*	0,51*
LRE	0,81*	0,36*	0,87*	0,08	0,25	-0,70*	0,74*	-0,42*
LGZE	-0,76*	-0,18	-0,96*	-0,02	-0,32*	-0,99*	0,18	-0,89*
HGZE	0,27	0,12	-0,48*	-0,31*	0,30*	0,99*	-0,06	0,95*
RLNU	1,00*	0,34*	1,00*	0,04	0,98*	0,48*	0,96*	0,24

Tableau 3.8 – Coefficients de corrélation de Spearman entre les index de texture (calculés soit après une discrétisation relative ou absolue) et VM ou SUVmax dans les tumeurs pulmonaires (VOI-T) et le tissu sain (VOI-F). * signifie que le coefficient de corrélation est significatif au seuil de 5%.

Relation entre les index de texture calculés avec la discrétisation absolue et le VM ou SUVmax

La corrélation entre les paramètres de texture calculés après une discrétisation absolue et VM est plus faible (cf. Figure 3.7) par rapport aux mêmes index calculés après une discrétisation relative. Les coefficients de corrélation de Spearman ($|r_s|$) varient entre 0,02 et 0,75, sauf pour RLNU ($r_s=0,98$ dans VOI-T et 0,96 dans VOI-F, cf. Tableau 3.8).

Néanmoins, la discrétisation absolue entraîne une corrélation plus importante entre les index de texture et SUVmax (cf. Figure 3.8) avec des coefficients de corrélation ($|r_s|$) entre 0,24 et 0,99, sauf pour le fantôme où le SUVmax est constant et où la corrélation avec SUVmax ne peut donc pas être étudiée.

3.4.3.3 Discussion

Le premier objectif de cette étude était de démontrer le lien entre les index de texture et le VM ou SUVmax pour des paramètres utilisés en analyse de texture des images TEP et avec une méthode de discrétisation relative. À la différence de [Brooks and Grigsby, 2014], nous avons utilisé un fantôme rempli uniformément de 18F-FDG pour notre étude. Dans ce fantôme, la texture mesurée à partir de la distribution de 18F-FDG vient donc uniquement de celle des images TEP qui est spatialement corrélée dans les images reconstruites [Buvat, 2002] et non d'une texture « physiologique ». Le comportement des index de texture en fonction du volume utilisé pour les calculer nous a permis d'étudier l'index sans être affecté par un signal physiologique sous-jacent. Nous avons montré que l'évolution de Entropy en fonction du VM (cf. Figure 3.5) est très similaire à ce qui avait été précédemment rapportée dans [Brooks and Grigsby, 2014] démontrant la pertinence de notre étude sur fantôme pour caractériser la relation entre les paramètres de texture et le VM. De plus, nous avons montré que d'autres paramètres se comportaient comme Entropy vis-à-vis de leur dépendance au VM, avec une forte corrélation pour les VOI de petites tailles et une stabilisation à partir d'un certain volume. Cependant, le volume nécessaire pour atteindre le plateau varie d'un index à l'autre (cf. Figure 3.5). De plus, certains paramètres n'atteignent jamais le plateau, confirmant que ces index sont de médiocres descripteurs de la texture, comme RLNU que nous avons identifié précédemment comme appartenant au groupe de corrélation incluant le VM. Nous avons également mis en évidence que la relation entre les index de texture et le VM est similaire pour le fantôme et les données cliniques, démontrant pour

la première fois que l'approche sur fantôme est adéquate pour caractériser la dépendance des index au VM comme observée chez les patients, malgré des différences dans la distribution des SUV et du rapport signal-sur-bruit (fantôme : 11 kBq/ml pour un temps d'acquisition de 3 min ; patients : 3,0–3,5 kBq/ml pour 1,45 min d'acquisition par position de lit).

Dans leurs conclusions, [Brooks and Grigsby, 2014] recommandent de calculer l'index Entropy seulement pour des régions de volume supérieur à 45 ml, ce qui limite sensiblement son utilisation en clinique. L'utilisation de la discrétisation absolue pour le calcul des paramètres de texture supprime cette limitation et permet le calcul des index sans introduire de biais associé au volume tumoral. En revanche, la discrétisation absolue introduit une dépendance des index de texture au SUV (cf. Figure 3.8), qui semble cependant plus justifiée. En effet, il est probable que si la région tumorale inclut une large gamme de SUV, la tumeur sera plus hétérogène que si tous les SUV sont similaires. Par conséquent, supprimer l'information de l'intervalle des SUV en utilisant l'équation 3.2 pour l'étape de discrétisation n'apparaît pas judicieux pour évaluer la texture tumorale. Comme l'illustre la Figure 3.3 obtenue à partir des données fantôme, la discrétisation relative crée artificiellement de la texture dans une zone qui initialement était homogène (sans structure sous-jacente) avec de très faibles variations de SUV, entre 0,8 et 1,2 unités SUV ici. À l'opposé, la discrétisation absolue permet bien d'obtenir une image homogène du fantôme en occupant principalement un seul niveau après discrétisation. Le fait que les index soient corrélés avec SUVmax (cf. Figure 3.8) ne semble donc pas contre-intuitif puisque l'intervalle de SUV reflète l'hétérogénéité tumorale.

La corrélation obtenue entre certains index de texture calculés avec la méthode relative et le VM correspondrait donc à un artefact de calcul et non à une mesure de l'hétérogénéité tumorale. À l'inverse, la corrélation obtenue entre les index de texture calculés avec la méthode absolue et SUVmax est une corrélation plus « physiologique » liée au processus tumoral. Cette interprétation est confortée par le fait qu'on la retrouve plus fortement marquée dans les VOI tumorales que dans le tissu sain (cf. Tableau 3.8).

Discrétiser entre deux bornes fixes sur l'ensemble des VOI préserve une information utile des SUV pour le calcul des index. Nous avons choisi arbitrairement 0 unité SUV comme borne basse et 20 unités SUV comme borne haute, ce qui correspond à l'intervalle de variations des SUV typiquement observé en clinique. Pour évaluer l'impact du choix arbitraire de la borne haute, nous avons étudié la robustesse des résultats pour l'évaluation clinique dans le chapitre suivant en fixant cette borne égale à 15 ou 25 unités SUV.

La méthode de discrétisation absolue est proche de celle proposée par [Leijenaar et al., 2013], où les auteurs discrétisent les intensités des voxels avec une largeur de classe (ou *bin* en anglais) fixée à 0,5 unités SUV. Cette approche équivaut à utiliser une discrétisation absolue avec une borne haute égale à 32 unités SUV. Avec une borne haute de 20 unités SUV et 64 valeurs discrètes, la largeur de classe est approximativement de 0,3 unités SUV, ce qui rend les index plus sensibles à de petites variations de SUV par rapport à une largeur de classe de 0,5.

Avec la discrétisation absolue, nous avons souhaité vérifier les résultats obtenus pour la formation des groupes de corrélation ainsi que la robustesse vis-à-vis de la méthode de segmentation. À partir de la cohorte CPNPC-2, nous avons comparé la moyenne des r obtenus dans les groupes de corrélation identifiés précédemment à partir des deux méthodes de discrétisation (cf. Tableau 3.9). Nos résultats montrent que les groupes restent globalement cohérents avec trois réorganisations : Correlation et Coarseness rejoignent le groupe n°7 et LZHG forme un nouveau groupe à part. Les groupes n°3, 5 et 6 ainsi que Busyness et LZLGE restent inchangés. La comparaison de la sensibilité des index à la méthode de segmentation montre (cf. Tableau 3.5 vs. Tableau 3.10) que les conclusions restent

également similaires. Seize index restent dans la même catégorie d'écart moyen, 6 ont un écart moyen qui diminue alors que pour 9 index, il augmente avec la discrétisation absolue par rapport à la discrétisation relative. Par exemple, LZHG conduit à un écart moyen important (25%) quand il est calculé avec la méthode absolue comparé à la méthode relative (de 3,5 à 6,5%), et cette instabilité peut expliquer le changement de groupe de cet index. Pour la variabilité, 29 index restent dans la même classe de variabilité (5-Faible, 10-Modérée, 4-Forte), 2 index passent d'une variabilité Modérée à Faible et 10 d'une variabilité Forte à Modérée.

Pour le choix des index à calculer, il apparaît donc nécessaire de :

- choisir un index par groupe ;
- sélectionner un index n'ayant pas changé de groupe entre les deux types de discrétisation pour qu'il soit utilisable à la fois après DR ou DA ;
- préférer les index les plus stables et les moins variables suivant la méthode de segmentation.

Au vu de ces critères, Homogeneity (groupe n°1), Entropy (groupe n°2), SRE (groupe n°3), et LRE (groupe n°4) apparaissent les plus robustes à la segmentation dans leur groupe respectif (en gras dans le Tableau 3.9). Pour les groupes n°5 et 6, LGZE et HGZE apparaissent parmi les moins sensibles de leur groupe à la segmentation (avec DR ou DA) et présentent l'avantage de porter uniquement une information sur les niveaux de gris sans être combinés avec la taille des zones ou des runs (comme LRLGE ou LRLHGE par exemple). Pour le groupe n°7, aucun index ne sera retenu du fait de leur forte corrélation avec le VM pour les deux méthodes de discrétisation. Enfin, les index Busyness et LZLGE se seront pas utilisés dans la suite compte tenu de leur manque de robustesse.

Groupe	Index	DR	DA	DA avec ajustement des groupes
1	Homogeneity -Correlation-Contrast(CO) -Dissimilarity-Contrast(NGLDM)	0,92±0,05	0,57±0,41	sans Correlation → 0,88±0,09
2	Energy- Entropy -Coarseness	0,91±0,4	0,67±0,19	sans Coarseness → 0,81
3	SRE -RP-SZE-ZP	0,97±0,03	0,86±0,09	-
4	LRE -LZE-LZHGE	0,84±0,09	0,50±0,31	sans LZHG → 0,84
5	LGRE-SRLGE-LRLGE- LGZE -SZLGE	0,92±0,08	0,92±0,08	-
6	HGRE-SRHGE-LRHGE- HGZE -SZHGE	0,89±0,11	0,96±0,05	-
7	GLNUr-RLNU-GLNUz-ZLNU-VM-TLG	0,94±0,07	0,81±0,16	avec Correlation et Coarseness → 0,75±0,15
9	Busyness	-	-	-
10	LZLGE	-	-	-

Tableau 3.9 – Moyenne et écart-type de r sur chaque paire d'index présents dans les groupes de corrélation en fonction de la discrétisation utilisée : relative (DR) ou absolue (DA). En gras, les 6 index de texture sélectionnés pour leur robustesse vis-à-vis de la méthode de segmentation.

Groupe	Index	CPNPC-2	Ecart moyen	Varibilité
1	Homogeneity	2,7 ± 5,7	+	Faible
	Contrast(CO)	-9,6 ± 15,7	+	Modérée
	Dissimilarity	-5,3 ± 9,3	+	Faible
	Contrast(NGLDM)	-10,5 ± 25,4	++	Modérée
2	Energy	16,9 ± 29,2	++	Modérée
	Entropy	-2,9 ± 8,1	+	Faible
3	SRE	-0,9 ± 3,2	+	Faible
	RP	-1,3 ± 4,4	+	Faible
	SZE	-0,0 ± 14,4	+	Faible
	ZP	-8,1 ± 27,3	+	Modérée
4	LRE	3,7 ± 12,2	+	Faible
	LZE	19,9 ± 51,8	++	Forte
5	LGRE	-11,1 ± 32,3	++	Modérée
	SRLGE	-12,0 ± 30,1	++	Modérée
	LRLGE	-7,9 ± 40,5	+	Modérée
	LGZE	-8,7 ± 34,8	+	Modérée
	SZLGE	-6,5 ± 40,2	+	Modérée
6	HGRE	5,0 ± 22,7	+	Modérée
	SRHGE	7,9 ± 17,7	+	Modérée
	LRHGE	4,4 ± 24,3	+	Modérée
	HGZE	5,7 ± 20,1	+	Modérée
	SZHGE	4,4 ± 24,2	+	Modérée
7	GLNUr	2,5 ± 40,3	+	Modérée
	RLNU	-4,6 ± 44,5	+	Modérée
	GLNUz	-4,3 ± 24,9	+	Modérée
	ZLNU	-9,7 ± 33,6	+	Modérée
	VM	-2,8 ± 48,6	+	Modérée
	TLG	1,1 ± 38,0	+	Modérée
	Correlation	-49,0 ± 231,1	+++	Forte
	Coarseness	1,9 ± 41,7	+	Modérée
8	SUVmax	0 ± 0	+	Faible
	SUVmean	2,9 ± 14,8	+	Faible
	SUVpeak	0 ± 0	+	Faible
9	Busyness	72,3 ± 926,5	+++	Forte
10	LZLGE	6,2 ± 70,7	+	Forte
11	LZHGE	25,4 ± 40,3	+++	Modérée

Tableau 3.10 – Moyennes (M_δ) et Ecart-types (sd_δ) de la différence relative entre un seuillage fixe et un seuillage adaptatif pour la cohorte CPNPC-2 avec la discrétisation absolue.

Trois catégories d'écart moyen ont été définies : + signifie $M_\delta(index) < 10\%$, ++ signifie $10\% \leq M_\delta(index) \leq 20\%$ et +++ signifie $M_\delta(index) > 20\%$.

Trois catégories de variabilité ont été définies : Faible signifie $sd_\delta(index) \leq sd_\delta(SUVmean)$, Modérée $sd_\delta(SUVmean) < sd_\delta(index) \leq sd_\delta(VM)$ et Forte $sd_\delta(VM) < sd_\delta(index)$.

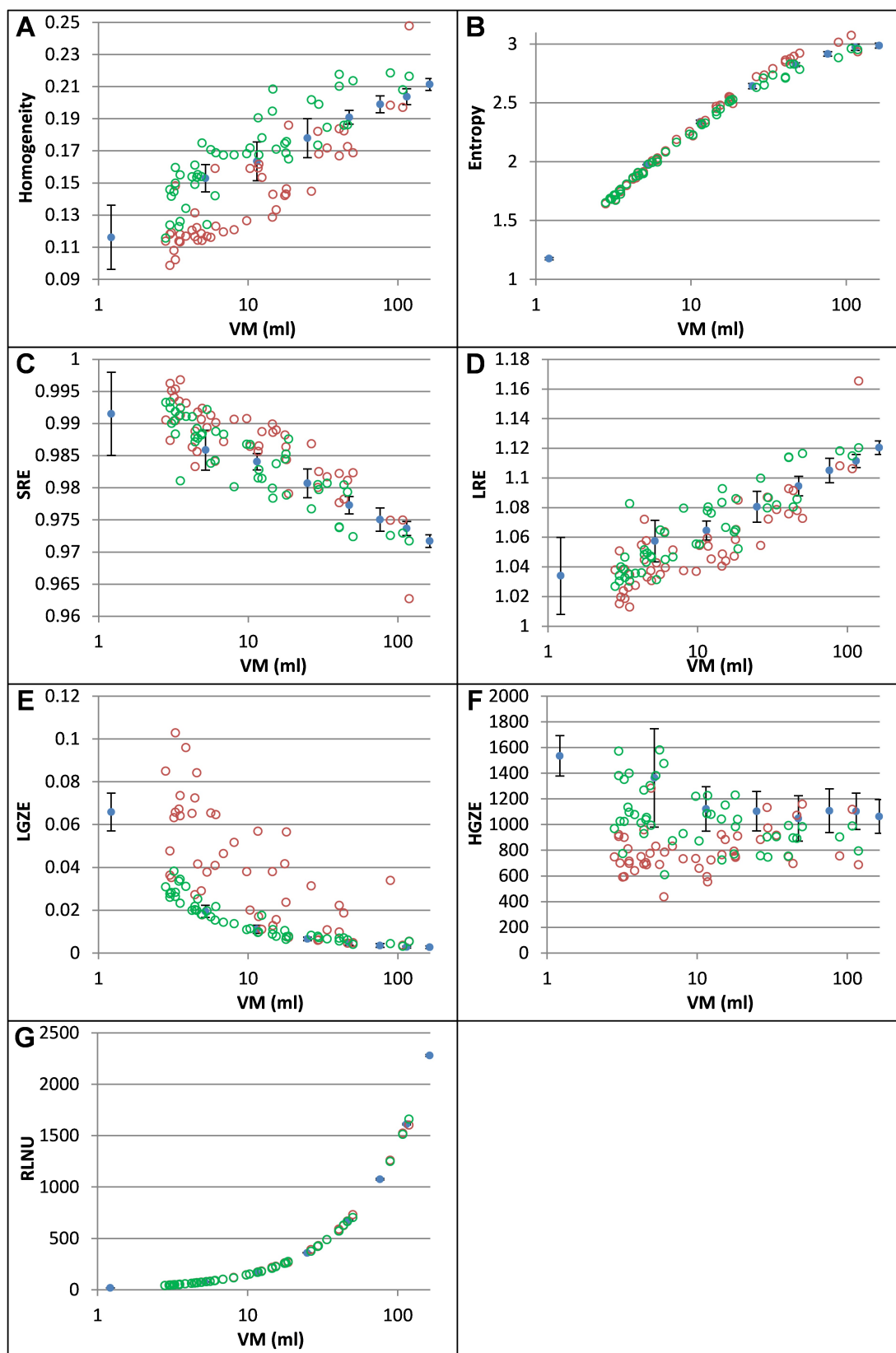


FIGURE 3.5 – Graphiques de Homogeneity (A), Entropy (B), SRE (C), LRE (D), LGZE (E), HGZE (F) et RLNU (G) obtenus après discrétisation relative en fonction du VM pour le fantôme (bleu), les tumeurs pulmonaires (rouge) et le tissu sain (vert). Les barres d'erreurs correspondent à un écart-type sur les 5 positions des sphères dans le fantôme.

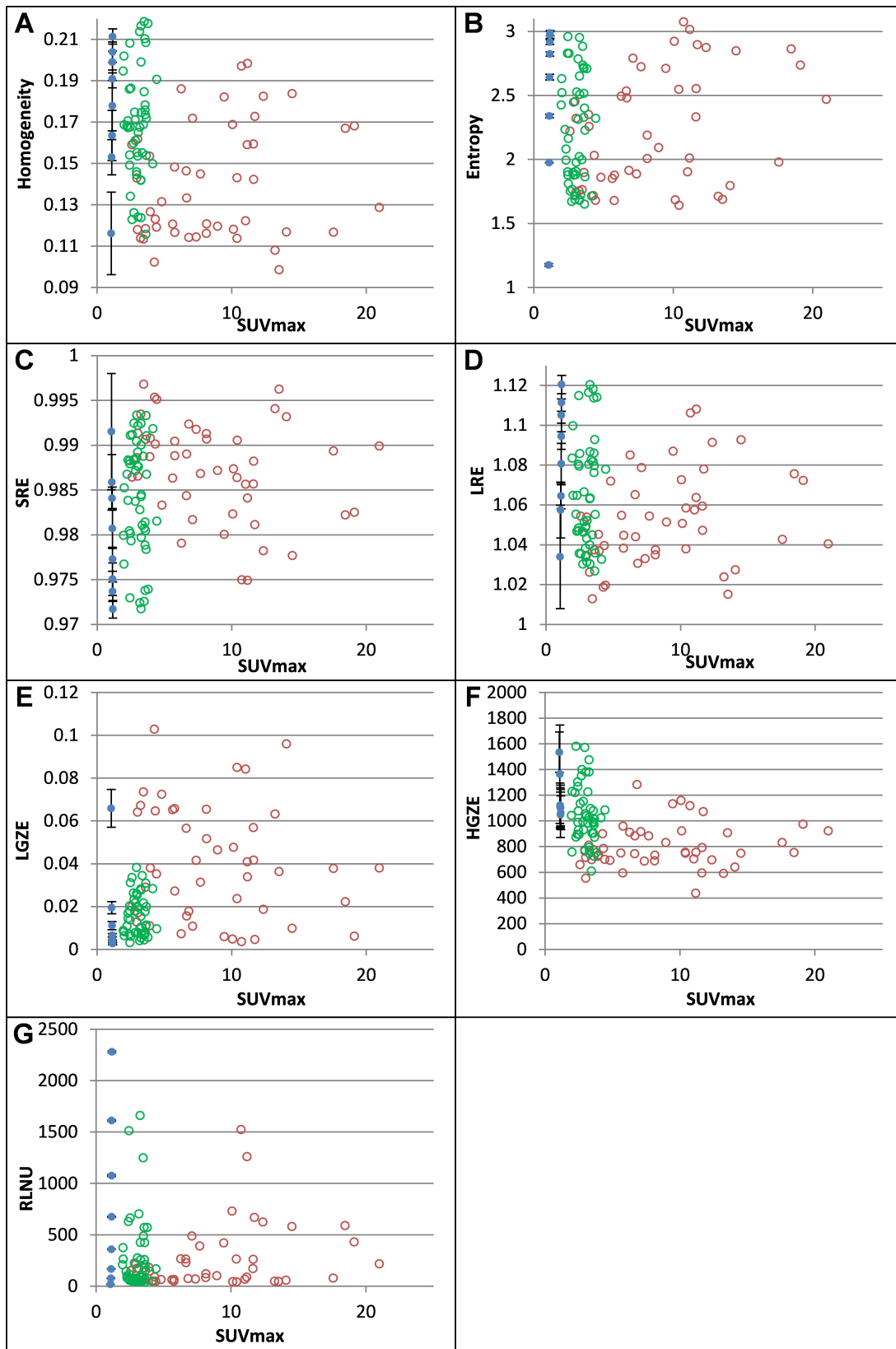


FIGURE 3.6 – Graphiques de Homogeneity (A), Entropy (B), SRE (C), LRE (D), LGZE (E), HGZE (F) et RLNU (G) obtenus après discrétisation relative en fonction du SUVmax pour le fantôme (bleu), les tumeurs pulmonaires (rouge) et le tissu sain (vert). Les barres d'erreurs correspondent à un écart-type sur les 5 positions des sphères dans le fantôme.

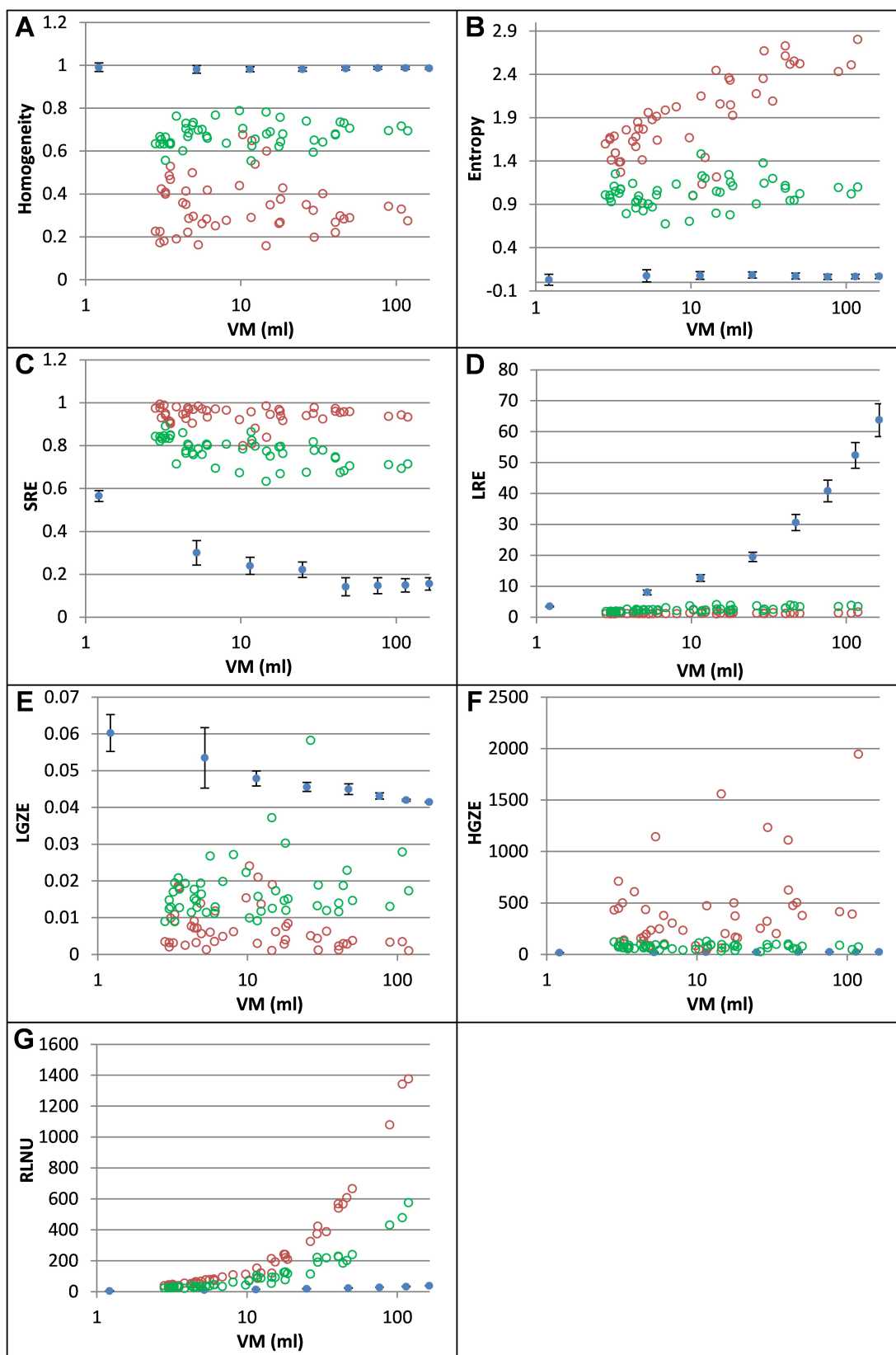


FIGURE 3.7 – Graphiques de Homogeneity (A), Entropy (B), SRE (C), LRE (D), LGZE (E), HGZE (F) et RLNU (G) obtenus après discrétisation absolue en fonction du VM pour le fantôme (bleu), les tumeurs pulmonaires (rouge) et le tissu sain (vert). Les barres d'erreurs correspondent à un écart-type sur les 5 positions des sphères dans le fantôme.

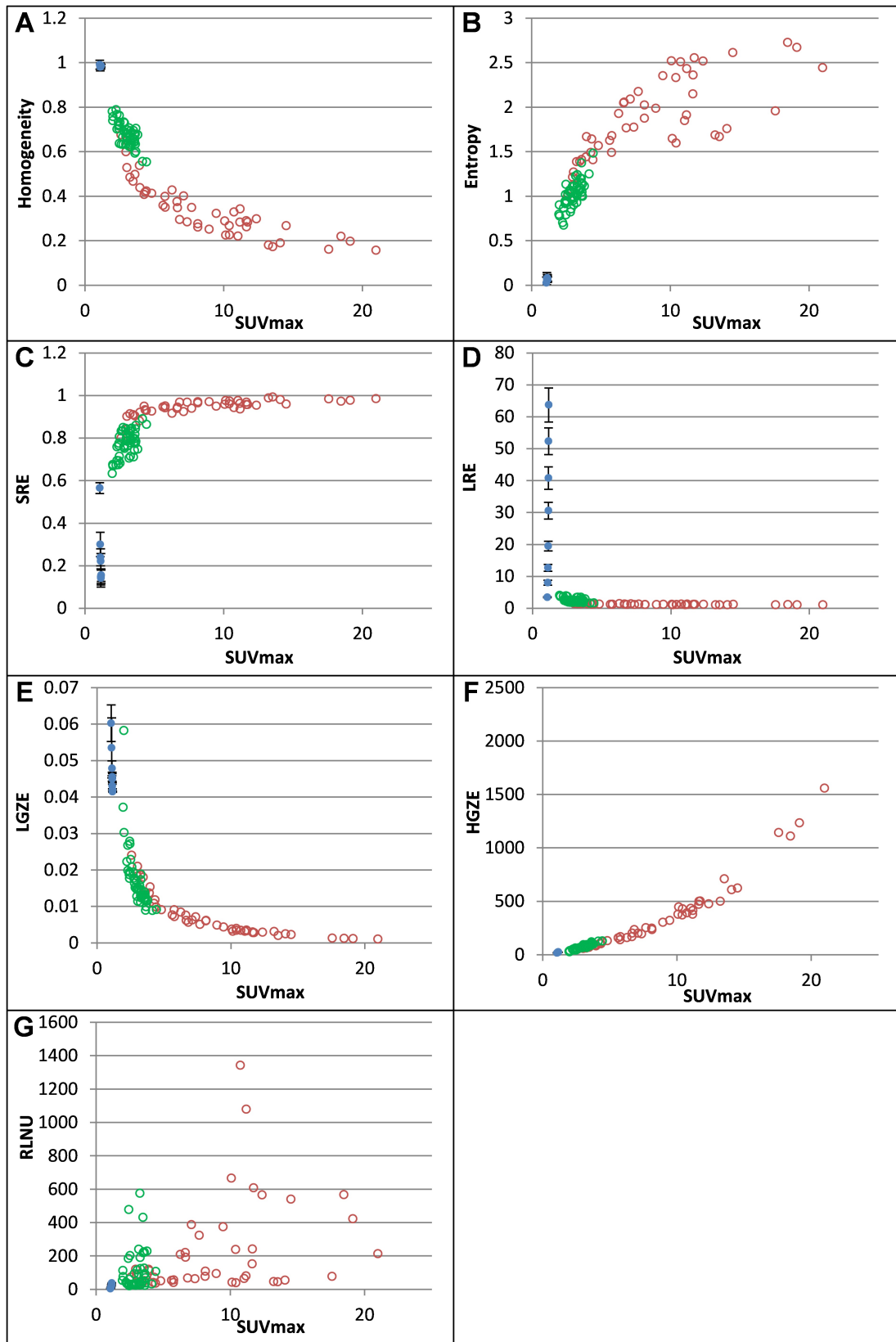


FIGURE 3.8 – Graphiques de Homogeneity (A), Entropy (B), SRE (C), LRE (D), LGZE (E), HGZE (F) et RLNU (G) obtenus après discrétisation absolue en fonction du SUVmax pour le fantôme (bleu), les tumeurs pulmonaires (rouge) et le tissu sain (vert). Les barres d'erreurs correspondent à un écart-type sur les 5 positions des sphères dans le fantôme.

3.5 Interprétation des paramètres de texture en TEP

La compréhension des étapes du calcul influençant les paramètres de texture a été indispensable pour calculer au mieux les index. Toutefois, pour être utilisé en clinique, il est indispensable de comprendre la signification des index de texture et quelles valeurs prennent les paramètres suivant la distribution spatiale de la fixation du traceur à l'intérieur du volume tumoral. C'est le but de l'étude présentée maintenant, qui a consisté à évaluer la variation des index suivant différents types de fixation, afin de fournir des points de repères pour l'interprétation de ces paramètres.

3.5.1 Matériels & méthodes

Pour décrire la variation des index de texture suivant différents types d'hétérogénéité de fixation, nous avons créé quinze simulations simples de sphères. La Figure 3.9 illustre les grandes étapes de la construction de ces simulations. Une taille de voxels de 4 mm de côté a été choisie arbitrairement. D'abord, nous avons simulé un fond uniforme de fixation 1 unité SUV, affecté par un bruit gaussien d'écart-type égal à 1. Dans ce fond, nous avons ensuite inséré une sphère de rayon 14 voxels avec une fixation égale à 8 unités SUV. Enfin, pour imiter grossièrement la résolution spatiale limitée des images TEP, nous avons lissé les images avec un filtre gaussien afin de modéliser une résolution des images TEP de 7 mm. Nous avons analysé les sphères résultantes en utilisant deux VOI : le premier correspond à une sphère de 12 voxels de rayon afin que les bords de la sphère soient exclus, alors que le second VOI correspond à une sphère de 14 voxels où les bords sont donc inclus.

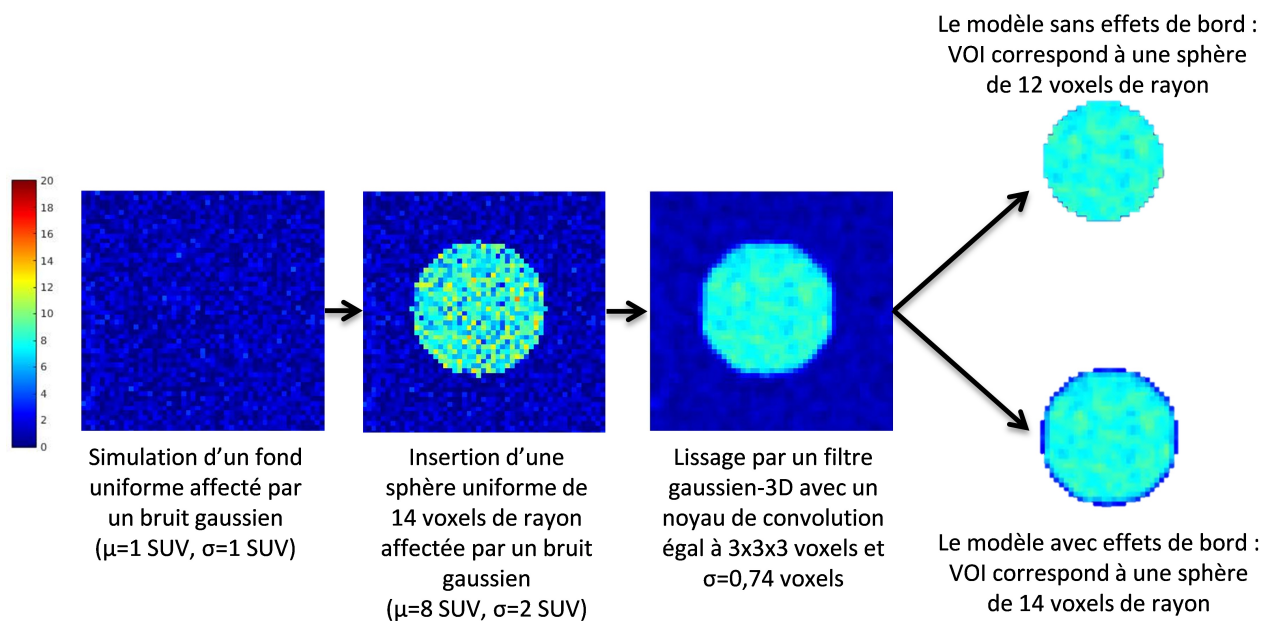


FIGURE 3.9 – Étapes de création des modèles de sphère avec et sans effets de bord. (μ : moyenne de la distribution, σ : écart-type).

Pour imiter l'hétérogénéité de fixation dans les tumeurs, nous avons créé des variations de ces deux sphères, comme illustré sur la Figure 3.10. Les sphères #1, #8 et #9 représentent des distributions homogènes mais avec différents niveaux de fixation et/ou de variabilité. Les sphères #2–7 et #10–15 correspondent à des « tumeurs » avec des hétérogénéités métaboliques internes, soit des sous-régions hypométaboliques modélisant grossièrement des nécroses, soit des sous-régions hypermétaboliques

qui pourraient correspondre à des clones avides de FDG dans des images TEP au 18F-FDG ou à des régions hypoxiques quand les images TEP sont obtenues avec un traceur ciblant l'hypoxie. Tous ces modèles ont été étudiés avec et sans effets de bord. Enfin, pour caractériser la variabilité des mesures, les 15 modèles ont été simulés 20 fois chacun.

Une fois les modèles construits, sur chaque VOI, nous avons calculé le SUVmax, SUVmean et le VM ainsi qu'un index par groupe de corrélation : Homogeneity, Entropy, SRE, LRE, LGZE et HGZE (comme décrit précédemment), après une discrétisation absolue avec 64 niveaux (entre 0 et 20 unités SUV). Nous avons ensuite tracé la variation de ces index pour les quinze modèles simulés. Les caractéristiques du calcul des index de texture sont résumées dans le tableau ci-dessous.

Cohorte(s)	Segmentation	Taux (D)	Formule
Simulations	R=14 et R=12 voxels	64	DA

3.5.2 Résultats

Les graphiques des index de texture, SUVmax et SUVmean en fonction des modèles simulés sans effets de bord (cf. Figure 3.11) montrent les valeurs des différents index et leur capacité à refléter différents types d'hétérogénéité. La Figure 3.11 montre 4 types de variation pour les index de texture. Homogeneity et LRE, d'une part, et Entropy et SRE, d'autre part, présentent les mêmes variations suivant les modèles. LGZE et HGZE varient en directions opposées et différemment des quatre autres index mentionnés précédemment.

Homogeneity, LRE, Entropy et SRE identifient la sphère #1 comme la plus homogène parmi les sphères simulées avec une valeur d'index toujours supérieure ou inférieure à celle mesurée dans les autres sphères, à la différence de HGZE, LGZE, SUVmax et SUVmean.

Quand des sphères internes sont ajoutées au modèle #1, Homogeneity et LRE décroissent et Entropy et SRE augmentent quelque soit la fixation des sphères internes, et cette décroissance ou croissance est proportionnelle au nombre ou à la taille des sphères insérées. Ces paramètres ne sont donc pas capables de distinguer des hypofixations des hyperfixations insérées dans la sphère. À l'inverse, HGZE et SUVmean sont sensibles au type d'hétérogénéité : ils diminuent quand des sphères hypofixantes sont ajoutées et augmentent quand ce sont des sphères hyperfixantes. Ces variations sont également influencées par la taille et le nombre de sphères ajoutées. LGZE varie de manière opposée comparé à HGZE. Comme attendu, SUVmax reste stable quand des sphères hypofixantes sont ajoutées et augmente quand ce sont des sphères hyperfixantes.

Concernant la distinction du lieu de la « nécrose », aucun paramètre ne semble capable de distinguer une hypofixation centrale (#10–12) d'une hypofixation périphérique (#13–15).

Quand on compare les sphères #1 et 8 qui diffèrent uniquement par leur activité moyenne (SUV=8 et 14) mais avec le même écart-type (2 unités SUV), la valeur de Homogeneity, LRE, Entropy et SRE reste stable, alors que HGZE, SUVmax et SUVmean augmentent fortement et LGZE diminue fortement.


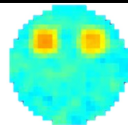
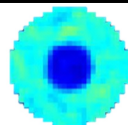
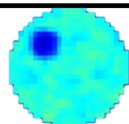
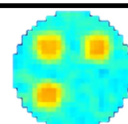
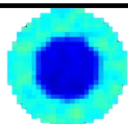

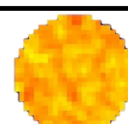
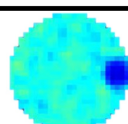
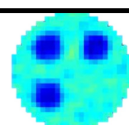
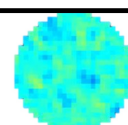
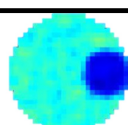
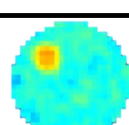
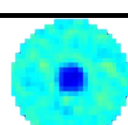
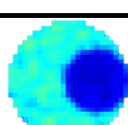
#1  Sphère uniforme de rayon 12 voxels affectée par un bruit gaussien ($\mu=8$ SUV, $\sigma=2$ SUV)	#6  Identique à la sphère #1 avec en plus deux sphères internes hyperfixantes ($\mu=14$ SUV, $\sigma=1$ SUV, $R=3$ voxels)	#11  Identique à la sphère #1 avec en plus une sphère interne hypofixante ($\mu=2$ SUV, $\sigma=1$ SUV, $R=5$ voxels)
#2  Identique à la sphère #1 avec en plus une sphère interne hypofixante ($\mu=2$ SUV, $\sigma=1$ SUV, $R=3$ voxels)	#7  Identique à la sphère #1 avec en plus trois sphères internes hyperfixantes ($\mu=14$ SUV, $\sigma=1$ SUV, $R=3$ voxels)	#12  Identique à la sphère #1 avec en plus une sphère interne hypofixante ($\mu=2$ SUV, $\sigma=1$ SUV, $R=7$ voxels)
#3  Identique à la sphère #1 avec en plus deux sphères internes hypofixantes ($\mu=2$ SUV, $\sigma=1$ SUV, $R=3$ voxels)	#8  Sphère uniforme de rayon 12 voxels affectée par un bruit gaussien ($\mu=14$ SUV, $\sigma=2$ SUV)	#13  Identique à la sphère #1 avec en plus une sphère interne hypofixante ($\mu=2$ SUV, $\sigma=1$ SUV, $R=3$ voxels)
#4  Identique à la sphère #1 avec en plus trois sphères internes hypofixantes ($\mu=2$ SUV, $\sigma=1$ SUV, $R=3$ voxels)	#9  Sphère uniforme de rayon 12 voxels affectée par un bruit gaussien ($\mu=8$ SUV, $\sigma=3,5$ SUV)	#14  Identique à la sphère #1 avec en plus une sphère interne hypofixante ($\mu=2$ SUV, $\sigma=1$ SUV, $R=5$ voxels)
#5  Identique à la sphère #1 avec en plus une sphère interne hyperfixante ($\mu=14$ SUV, $\sigma=1$ SUV, $R=3$ voxels)	#10  Identique à la sphère #1 avec en plus une sphère interne hypofixante ($\mu=2$ SUV, $\sigma=1$ SUV, $R=3$ voxels)	#15  Identique à la sphère #1 avec en plus une sphère interne hypofixante ($\mu=2$ SUV, $\sigma=1$ SUV, $R=7$ voxels)

FIGURE 3.10 – Illustration des 15 modèles simulés sans effets de bord. (μ : moyenne de la distribution, σ : écart-type de la distribution, R : rayon de la (ou des) sphère(s) interne(s)).

Inversement, pour une activité moyenne donnée (SUV=8), quand l'écart-type de la distribution augmente de 2,0 à 3,5 unités SUV (comparaison des sphères #1 et 9), HGZE, LGZE et SUVmean restent stables, Homogeneity et LRE diminuent alors que Entropy, SRE et SUVmax augmentent.

Quand on inclut les effets de bord dans les VOI, la valeur absolue des index de texture change significativement mais les variations des paramètres en fonction des sphères simulées restent similaires (cf. Figure 3.12). Seul SUVmax n'est pas affecté par l'inclusion ou non des effets de bord.

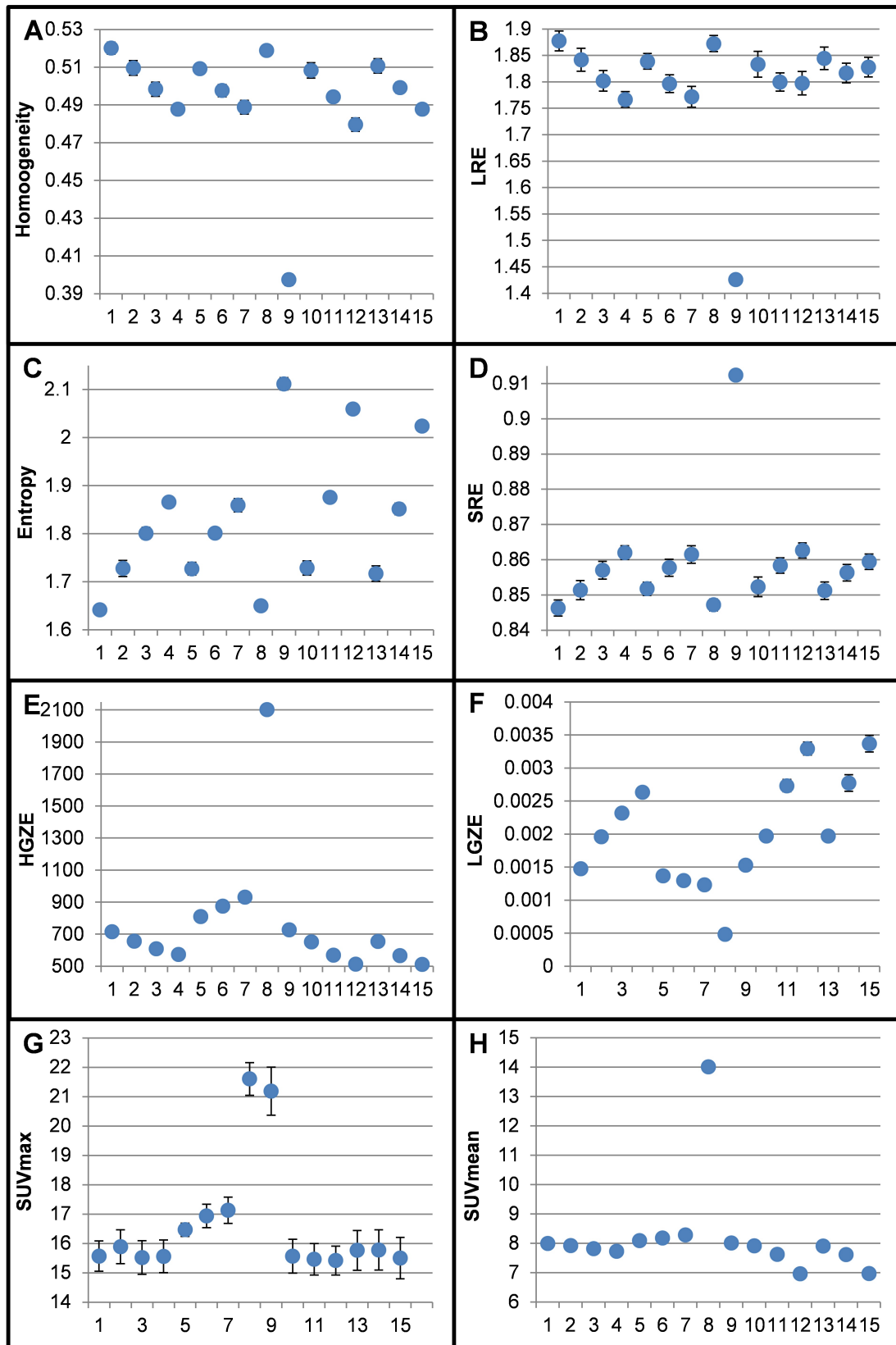


FIGURE 3.11 – Graphiques de Homogeneity (A), LRE (B), Entropy (C), SRE (D), HGZE (E), LGZE (F) pour les quinze modèles simulés sans effets de bord (moyenne et écart-type sur 20 simulations pour chaque sphère).

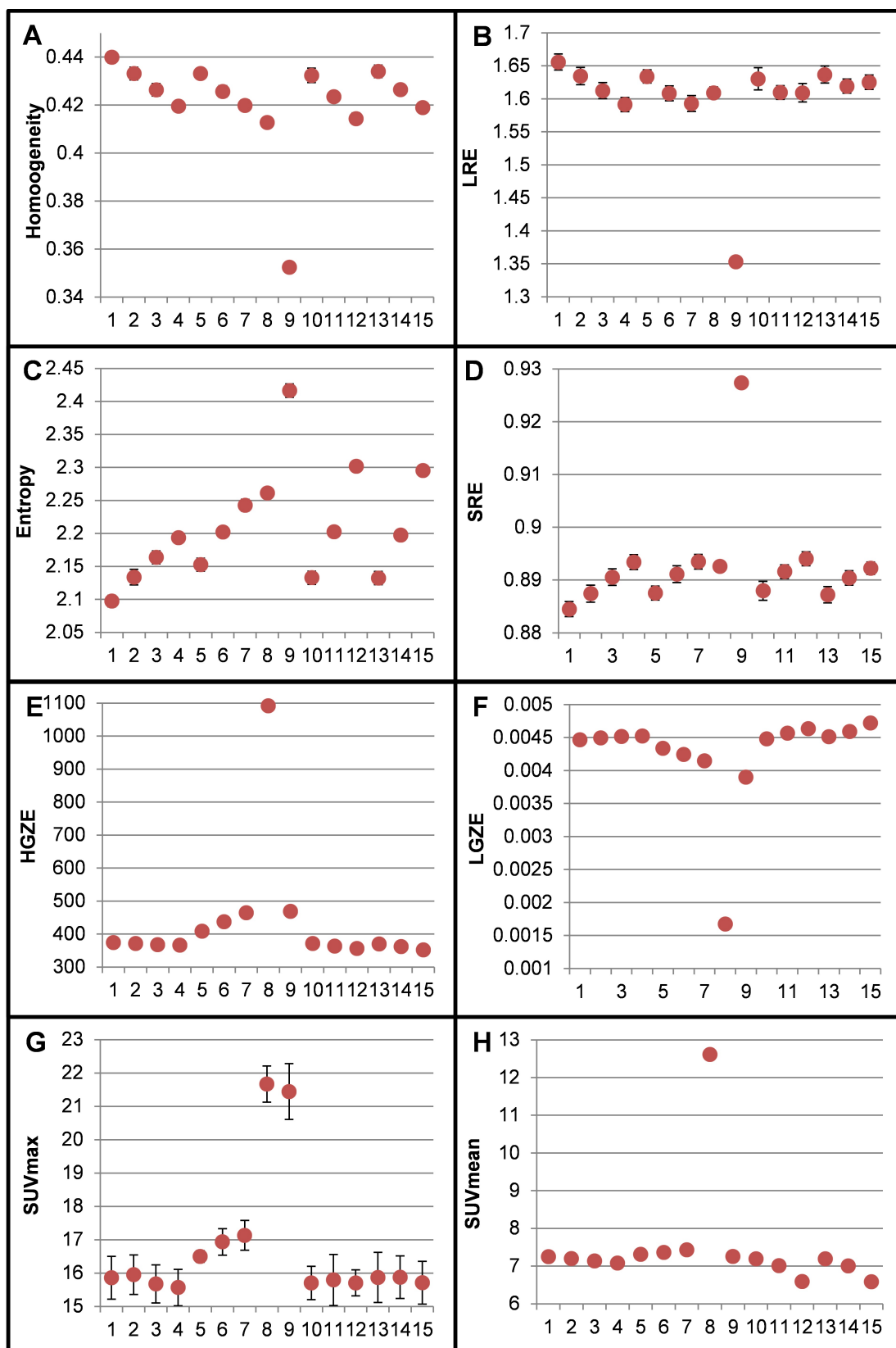


FIGURE 3.12 – Graphiques de Homogeneity (A), LRE (B), Entropy (C), SRE (D), HGZE (E), LGZE (F) pour les quinze modèles simulés avec effets de bord (moyenne et écart-type sur 20 simulations pour chaque sphère).

Les variations des index de texture et des SUV suivant les modèles simulés sont résumées dans le Tableau 3.11. Parmi les différents index testés, seul Homogeneity est borné entre 0 et 1. La sphère la plus homogène (#1) conduit à une valeur pour Homogeneity de $0,520 \pm 0,003$ (moyenne sur les 20 simulations ± 1 écart-type), $LRE = 1,88 \pm 0,02$, $Entropy = 1,64 \pm 0,01$, $SRE = 0,846 \pm 0,002$, $HGZE = 716 \pm 8$ et $LGZE = 0,00148 \pm 0,00002$ sans effets de bord (cf. Figure 3.11). Quand les bords sont inclus dans les calculs, la sphère #1 apparaît plus hétérogène avec un index $Homogeneity = 0,440 \pm 0,002$, $LRE = 1,66 \pm 0,01$, $Entropy = 2,10 \pm 0,01$, $SRE = 0,884 \pm 0,001$, $HGZE = 374 \pm 3$ et $LGZE = 0,00447 \pm 0,00006$ (cf. Figure 3.12).

	Homog. LRE	Entropy SRE	LGZE	HGZE	SUVmax	SUVmean
Ajout de sphères internes	↘	↗	Variable suivant la fixation	Variable suivant la fixation	Variable suivant la fixation	Variable suivant la fixation
Distinction entre hyperfixation et hypofixation	Non	Non	Hyper ↘ Hypo ↗	Hyper ↗ Hypo ↘	Hyper ↗ Hypo =	Hyper ↗ Hypo ↘
Augmentation de la fixation moyenne (μ) : #1 et #8	Non	Non	↘	↗	↗	↗
Augmentation de l'écart-type (σ) : #1 et #9	↘	↗	Non	Non	↗	Non
Localisation des hypofixations	Non	Non	Non	Non	Non	Non
Effets de bords	↘	↗	↗	↘	Non	↘

Tableau 3.11 – Résumé des variations des index de texture suivant la liste des critères testés grâce aux différentes simulations.

3.5.3 Discussion

Dans cette section, nous avons montré que les index de texture variaient suivant l'hétérogénéité de fixation simulée dans les sphères, permettant ainsi d'établir des points de repères pour l'interprétation des paramètres en clinique.

Tout d'abord, la distribution statistique des valeurs de voxels simulées concorde avec celle observée dans le fantôme comme le montre les tests de Kolmogorov-Smirnov réalisés entre la sphère #1 répétée 5 fois et 5 sphères segmentées automatiquement dans le fantôme (cf. Figure 3.16). En dépit du modèle assez simpliste de bruit suivi d'un lissage que nous avons utilisé, les simulations reproduisent donc correctement les variations de signal observées dans une région homogène lors d'une acquisition TEP réelle.

Ensuite, les variations des index de texture sont cohérentes avec l'analyse visuelle de l'hétérogénéité de fixation, à la différence de ce qui avait été rapporté dans [Tixier et al., 2014b] où l'index Homogeneity est plus élevé dans les tumeurs classées hétérogènes par deux experts (cf. Figure 3.13). Dans notre étude, les sphères les plus homogènes visuellement (sphères #1 et 8) sont bien celles qui ont un index Homogeneity le plus élevé. Comme attendu, Entropy, qui correspond à une mesure du chaos [Gebejes and Huertas, 2013], varie en direction opposée par rapport à Homogeneity, alors que dans [Tixier et al., 2011], ces deux index varient étonnamment dans le même sens (cf. Figure 3.14). La raison de ces différences dans les résultats provient de l'étape de normalisation utilisée pour le calcul des index de texture. Dans cette section, nous avons discrétisé les intensités avec la méthode absolue alors que [Tixier et al., 2011, Tixier et al., 2014b] utilisent la méthode relative. Avec la discrétisation relative (cf. Figure 3.15), on constate que la valeur de Homogeneity, LRE, Entropy et SRE des sphères les plus homogènes (#1,8) apparaît être contre-intuitive par rapport aux valeurs observées dans les autres sphères (#2–7 et #9–15). Par exemple, avec la discrétisation relative, HGZE diminue quand les sphères internes sont hypofixantes, ce qui est incohérent puisque HGZE est supposé augmenter quand il y a plus de régions avec des hauts niveaux de gris (c'est-à-dire des hautes valeurs de SUV dans notre cas).

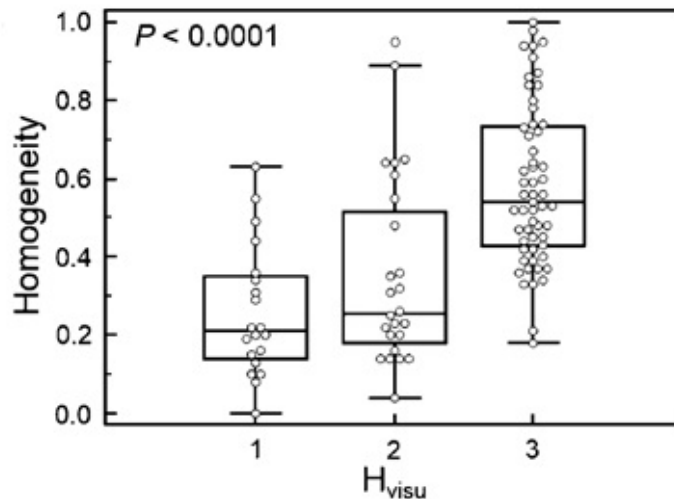


FIGURE 3.13 – Illustration extraite de [Tixier et al., 2014b]. Graphique de Homogeneity en fonction de trois niveaux d'hétérogénéité évalués visuellement par deux experts : $H_{visu} = 1$: la tumeur est visuellement homogène, $H_{visu} = 2$: la lésion est classée modérément hétérogène, $H_{visu} = 3$: la lésion est très hétérogène. L'index Homogeneity varie à l'opposé de la classification des experts.

Nos résultats suggèrent également que différents index de texture mesurent différents types d'hétérogénéité. En particulier, Homogeneity, LRE, Entropy et SRE ne sont pas sensibles au type de fixation ajoutée (hyposignal ou hypersignal), alors que HGZE et LGZE peuvent faire cette distinction. Ce constat est vérifié quelque soit la méthode de discrétisation (relative ou absolue).

Dans les images simulées, l'information de l'hétérogénéité capturée par LRE semble très similaire à celle capturée par Homogeneity, tout comme SRE, qui montre des variations très proches de celles de Entropy. Leijenaar et al. [Leijenaar et al., 2013] montrent que ces quatre paramètres sont très stables dans des études tests-retests, et ont une faible variabilité inter-observateur dans des cancers pulmonaires non à petites cellules, avec une plus forte robustesse de Homogeneity et Entropy

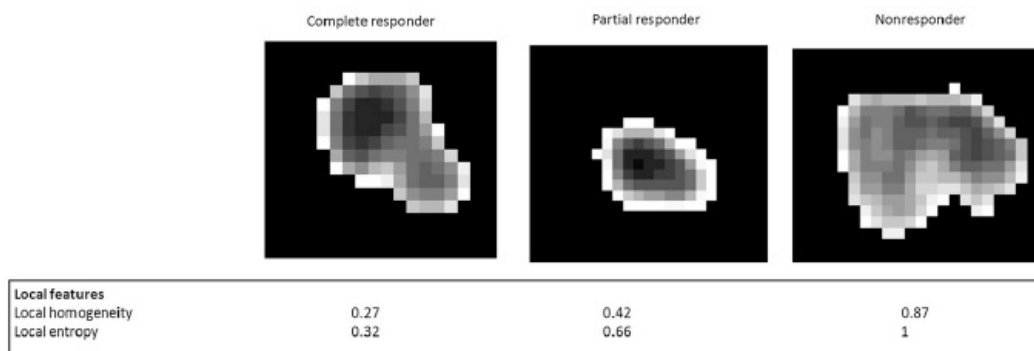


FIGURE 3.14 – Illustration extraite de [Tixier et al., 2011]. Valeurs de Homogeneity et Entropy pour trois lésions. Homogeneity et Entropy varient dans le même sens.

comparé à LRE et SRE. Homogeneity et Entropy présentent une meilleure reproductibilité que SUVmax et SUVmean dans le cancer de l'œsophage [Tixier et al., 2012]. Enfin, nous avons montré précédemment, tout comme [Hatt et al., 2013], que Homogeneity et Entropy sont peu ou moyennement influencés par la méthode de segmentation.

Par rapport à la sphère #1, les sphères #2 à 7 ou #10 à 15 peuvent être vues comme des modèles de « macro-hétérogénéités » où la taille des sphères incluses est comparable à la taille de la tumeur, alors que la sphère #9 peut être considérée comme un modèle de « micro-hétérogénéités » provenant d'une plus grande variabilité de la fixation (écart-type de la distribution) par rapport à la sphère #1. La Figure 3.11 suggère que LGZE et HGZE capturent mieux les macro-hétérogénéités que les micro-hétérogénéités alors que c'est l'inverse pour Homogeneity, LRE, Entropy et SRE. Dans le cancer de l'œsophage, Entropy est significativement corrélé au stade tumoral et ganglionnaire d'après [Dong et al., 2013]. Ce même index permet également de distinguer les tumeurs malignes dans les lésions osseuses et des tissus mous [Xu et al., 2014] et est associé à la survie globale dans des cancers pulmonaires non à petites cellules [Hatt et al., 2015]. El Naqa et al. [El Naqa et al., 2009] montrent que Homogeneity peut prédire la survie globale dans des cancers ORL et séparer le tissu mammaire sain du tissu mammaire tumoral [Willaime et al., 2013]. Nos résultats suggèrent donc que dans toutes ces études, si les conclusions sont confirmées avec un contrôle approprié des erreurs de type I [Chalkidou et al., 2015], les caractéristiques tumorales impliquées dans les tâches de classifications sont plutôt en rapport avec la micro-hétérogénéité qu'avec la macro-hétérogénéité.

Nos résultats détaillent également l'impact de l'inclusion des bords du VOI dans le calcul des index de texture. L'ajout des bords dans le VOI ne change pas la manière dont les index varient par rapport aux différentes sphères simulées (cf. Figures 3.11 et 3.12) mais affecte significativement la valeur absolue des paramètres de texture. Par exemple, pour la sphère #1, Homogeneity vaut (en moyenne) 0,52 sans effets de bord contre 0,44 avec effets de bords. Ainsi, il est difficile de comparer les valeurs des index de texture rapportées dans plusieurs articles puisqu'il est fort probable que la méthode utilisée pour déterminer le VOI diffère d'une publication à l'autre. Dans cette étude, l'inclusion des bords dans le VOI est plus importante que ce qui serait observé en pratique lors de l'utilisation de deux algorithmes de segmentation différents. Cependant, ces résultats mettent clairement en évidence le besoin d'une description précise de tous les paramètres qui peuvent influencer la valeur des index de texture afin de pouvoir comparer les valeurs absolues avant la réalisation d'études prospectives.

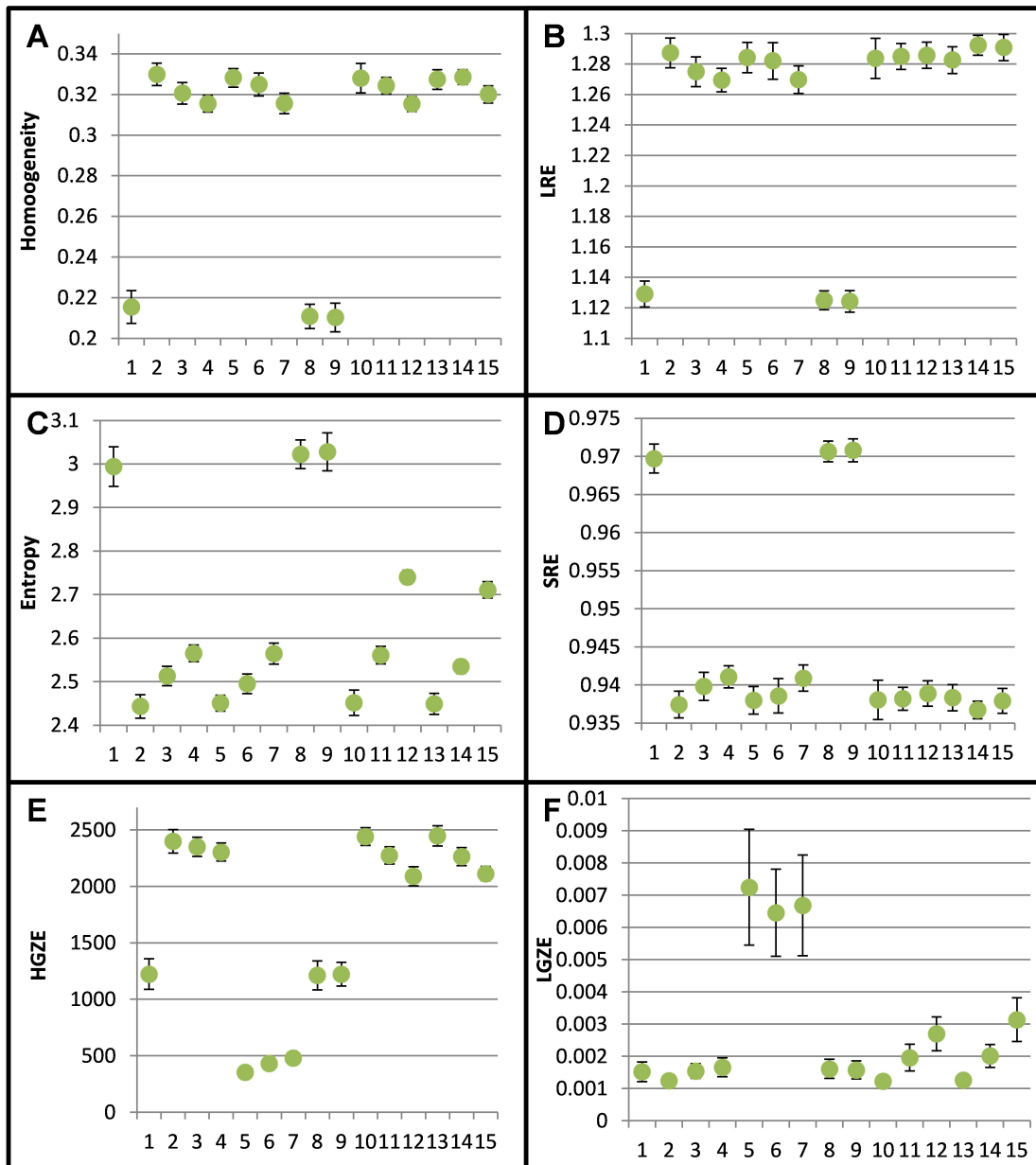


FIGURE 3.15 – Graphiques de Homogeneity (A), LRE (B), Entropy (C), SRE (D), HGZE (E), LGZE (F) pour les quinze modèles simulés avec effets de bord (moyenne et écart-type sur 20 simulations pour chaque sphère).

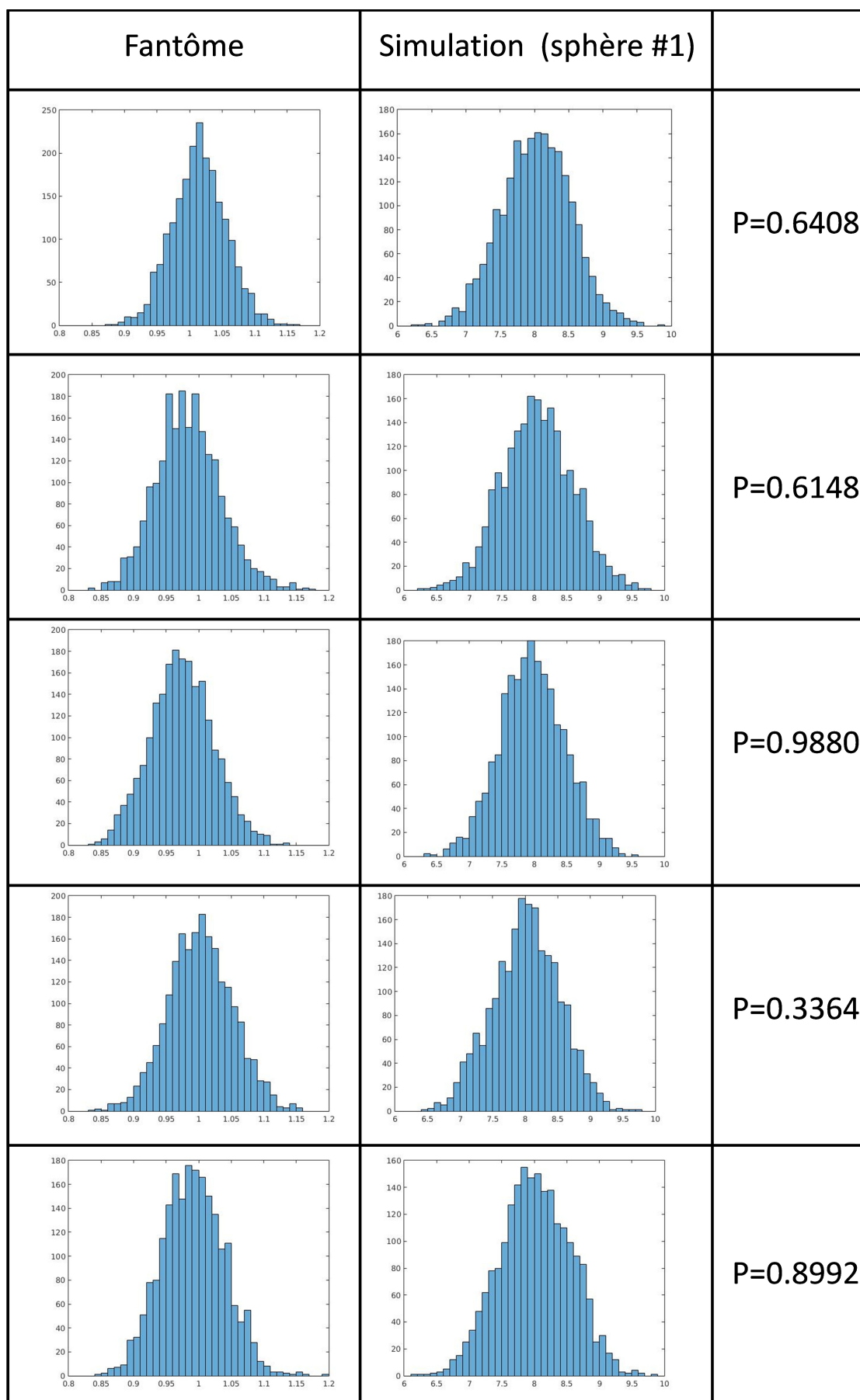


FIGURE 3.16 – Comparaison des histogrammes issus de VOI-S dans le fantôme et de simulations de la sphère #1, avec la valeur de p-value du test de Kolmogorov-Smirnov associé.

Au cours de ce chapitre, nous avons mené une étude méthodologique sur les paramètres de texture qui nous a conduits à fixer un protocole pour le calcul de ces index et nous a permis d'établir des points de repères pour leur interprétation en clinique. Pour une utilisation fiable de l'analyse de texture, nous recommandons donc :

- de segmenter les lésions avec une méthode reproductible (automatique ou semi-automatique) et non un contourage manuel ;
- de discrétiser avec au minimum 32 valeurs discrètes les intensités des voxels. Par défaut, nous choisirons 64 niveaux de gris dans la suite (comme dans le plupart des études de la littérature) ;
- d'utiliser la méthode de discrétisation absolue et non relative afin de limiter la corrélation avec le VM provenant d'un artefact de calcul ;
- de choisir un index par groupe de corrélation pour réduire le nombre de paramètres à calculer et le nombre de tests à effectuer, afin de limiter l'apparition de résultats faux positifs [Chalkidou et al., 2015] ;
- de toujours comparer les résultats avec ceux des index conventionnels.

Nous avons identifié six index prometteurs que nous évaluerons dans un contexte clinique au chapitre suivant en appliquant les règles mentionnées ci-dessus :

- Homogeneity (matrice de co-occurrence) qui mesure l'homogénéité de la distribution des niveaux de gris dans l'image ;
- Entropy (matrice de co-occurrence) qui est sensible au désordre dans l'image ;
- SRE (matrice GLRLM) qui mesure la distribution des courtes séries ;
- LRE (matrice GLRLM) qui, à l'inverse de SRE, est sensible aux longues séries ;
- LGZE (matrice GLZLM) qui quantifie les zones de bas niveaux de gris ;
- HGZE (matrice GLZLM) qui, à l'opposé de LGZE, renseigne sur les zones de haut niveaux de gris.

Chapitre 4

Évaluation clinique des index de texture

Après avoir établi, au Chapitre 3, un protocole strict pour le calcul des index de texture, nous avons évalué leur capacité à différencier les types de tissus et vérifié s'ils étaient corrélés à l'analyse visuelle, avant d'étudier leur relation avec les caractéristiques tumorales par rapport aux cinq index conventionnels utilisés plus ou moins fréquemment en routine clinique. En effet, les études présentées dans la littérature s'intéressent principalement aux capacités des paramètres de texture à prédire la réponse aux traitements ou la survie du patient. Très peu d'études ont cherché à déterminer si ces index, issus des images TEP, étaient sensibles aux spécificités des lésions tumorales.

4.1 Distinction entre les tissus sains et tumoraux

Cette première partie concerne la capacité des index de texture à distinguer les tissus sains des tissus tumoraux. En effet, il semble pertinent de montrer que l'organisation tissulaire est reflétée par les images TEP avant d'utiliser l'analyse de texture pour évaluer la réponse ou la survie du patient.

4.1.1 Matériels & méthodes

À partir des cohortes CM et CPNPC-2 présentées au Chapitre 3, nous avons segmenté la lésion primitive de chaque patient avec un seuillage adaptatif (cf. section 3.1) et reporté cette VOI tumorale (VOI-T) dans le foie de chaque patient (VOI-F) et dans la région controlatérale (VOI-C) soit dans le sein opposé pour la cohorte CM, soit dans le poumon opposé pour CPNPC-2. À partir des VOI segmentés, nous avons calculé cinq index conventionnels : SUVmax, SUVmean, SUVpeak, VM et TLG. Après avoir ré-échantillonné avec 64 niveaux les intensités des voxels à partir d'une discrétisation relative (DR) et absolue (DA) entre 0 et 15, 20 ou 25 unités SUV, nous avons calculé six index de texture : Homogeneity, Entropy, SRE, LRE, LGZE et HGZE. En effet, parmi les 10 groupes d'index précédemment identifiés pour leur robustesse vis-à-vis de la méthode de segmentation utilisée, nous avons choisi de calculer six paramètres de texture, un par groupe, sauf pour le groupe n°7 (corrélation très forte avec VM), le groupe n°8 (mesures SUV sans index de texture) et les groupes n°9 et 10 dont les deux index sont très influencés par la méthode de segmentation.

Pour chaque groupe de patients et chaque méthode de discrétisation, nous avons testé la capacité des index à distinguer : VOI-T vs. VOI-F, VOI-T vs. VOI-C et VOI-F vs. VOI-C à l'aide d'un test de Wilcoxon. Sachant qu'un grand nombre de tests statistiques sont effectués, une correction de Holm [Holm, 1979] est ensuite appliquée. Cette correction, moins sévère que celle de Bonferroni, permet de contrôler la probabilité d'une conclusion faussement positive (rejet de l'hypothèse nulle à tort) en la fixant à une valeur α (typiquement $\alpha = 5\%$), en procédant de la façon suivante.

Les p-values des k tests sont classées par ordre croissant et la formule suivante est appliquée :

$$p_i \leq \frac{\alpha}{k + 1 - i} \quad (4.1)$$

où i correspond au numéro d'ordre du test ($1 \leq i \leq k$), p_i à la p-value du test i après classement, k au nombre de tests ($2 \leq k \leq \infty$). Tant que p_i vérifie cette condition, l'hypothèse nulle est rejetée. Sinon, l'hypothèse nulle du test i et de tous ceux qui sont après ne peuvent être rejetées (la procédure s'arrête).

Nous nous sommes intéressés ensuite à la différenciation entre les types de tumeurs. Enfin, nous avons analysé la cohérence des résultats quand la borne supérieure utilisée pour la discrétisation absolue est égale à 15 ou 25 unités SUV, plutôt que 20 unités SUV.

4.1.2 Résultats

Nous avons comparé la valeur des index de texture calculés à partir de DR et DA dans le tissu tumoral, le tissu hépatique et le tissu controlatéral. Les p-values du test de Wilcoxon sont résumées dans le Tableau 4.1 pour la cohorte CPNCP-2 et dans le Tableau 4.2 pour la cohorte CM. Homogeneity, LGZE et HGZE sont significativement différents entre VOI-T et VOI-F, mais aucune p-value des index calculés avec DR n'est plus faible que celles des mesures SUV. Ceci suggère qu'aucun index calculé avec DR n'est meilleur que les mesures SUV pour séparer tissus hépatiques et tumoraux. Ce constat peut être fait sur les deux cohortes ainsi que pour la comparaison VOI-T vs. VOI-C et VOI-F vs. VOI-C. Avec la méthode DA, toutes les p-values sont plus faibles qu'avec la méthode DR et significatives même après la correction de Holm, suggérant une meilleure différenciation des types de tissus. Homogeneity, Entropy, SRE et LRE avec DA conduisent à une p-value inférieure à celle des SUV pour la distinction VOI-T vs. VOI-F. Par exemple, Entropy mesuré après DR ne permettait pas de distinguer les différents tissus dans les deux groupes de patients, alors qu'avec DA toutes les p-values sont inférieures à $1,09 \times 10^{-7}$ et significatives même après la correction des comparaisons multiples.

Avec la méthode DA, le tissu tumoral conduit à des valeurs d'Homogeneity, LRE et LGZE plus faibles que pour le tissu sain (hépatique ou controlatéral) ; à l'inverse, les valeurs d'Entropy, SRE, HGZE et SUVmax sont plus élevées. Dans les tissus sains, le tissu hépatique apparaît plus hétérogène que le tissu controlatéral. Ces observations sont vérifiées dans les deux groupes de patients (cf. Figures 4.1 et 4.2).

Nous avons ensuite cherché à comparer la texture observée entre les deux cohortes (Tableau 4.3). Les index de texture, après la correction de Holm, ne distinguent pas de différence entre le tissu hépatique des deux groupes de patients. Concernant les deux types tumoraux, seul Entropy semble caractériser le type de tumeurs avec un index plus élevé dans les tumeurs mammaires que dans les tumeurs pulmonaires (cf. Figure 4.3), ainsi qu'un VM et un SUVmax plus élevés.

Comme nous l'avons présenté au Chapitre 3, les bornes utilisées pour la discrétisation absolue ont été choisies arbitrairement. Nous avons donc souhaité vérifier si cela avait un impact sur les résultats en comparant la capacité des index suivant que la borne haute est égale à 15, 20 ou 25 unités SUV. Les résultats pour la cohorte CPNPC-2 pour la distinction VOI-T vs. VOI-F sont présentés dans le Tableau 4.4. Quelque soit la valeur choisie pour la borne haute, on ne constate pas d'impact significatif sur les p-values des index de texture. Par contre, la Figure 4.4 montre que la valeur absolue des index est influencée par le choix des bornes de discrétisation.

Index	Discrétisation relative			Discrétisation absolue		
	VOI-T vs. VOI-F	VOI-T vs. VOI-C	VOI-F vs. VOI-C	VOI-T vs. VOI-F	VOI-T vs. VOI-C	VOI-F vs. VOI-C
Homogeneity	7,47e-06*	1,11e-07*	0,0818	< 2,2e-16*	< 2,2e-16*	2,81e-16*
Entropy	0,7510	0,4858	0,6484	< 2,2e-16*	< 2,2e-16*	5,46e-16*
SRE	0,0505	0,0015*	0,1215	< 2,2e-16*	< 2,2e-16*	< 2,2e-16*
LRE	0,0432	0,0014*	0,1146	< 2,2e-16*	< 2,2e-16*	< 2,2e-16*
LGZE	7,19e-06*	0,0001*	0,0942	1,94e-11*	< 2,2e-16*	< 2,2e-16*
HGZE	1,12e-08*	0,0132*	0,0086	2,06e-12*	< 2,2e-16*	< 2,2e-16*
SUVmax	3,56e-16*	< 2,2e-16*	< 2,2e-16*	3,56e-16*	< 2,2e-16*	< 2,2e-16*
SUVmean	3,87e-12*	< 2,2e-16*	< 2,2e-16*	3,87e-12*	< 2,2e-16*	< 2,2e-16*
SUVpeak	1,20e-15*	< 2,2e-16*	< 2,2e-16*	1,20e-15*	< 2,2e-16*	< 2,2e-16*
VM	1,0	1,0	1,0	1,0	1,0	1,0
TLG	0,0102	2,06e-12*	1,45e-08*	0,0102*	2,06e-12*	1,45e-08*

Tableau 4.1 – P-values du test de Wilcoxon entre le tissu tumoral, le tissu hépatique et le tissu controlatéral dans la cohorte CPNPC-2 avec la discrétisation relative et absolue (entre 0 et 20 unités SUV). Les p-values en gras sont significatives au seuil de 5%. * signifie que la p-value reste significative après la correction de Holm.

Index	Discrétisation relative			Discrétisation absolue		
	VOI-T vs. VOI-F	VOI-T vs. VOI-C	VOI-F vs. VOI-C	VOI-T vs. VOI-F	VOI-T vs. VOI-C	VOI-F vs. VOI-C
Homogeneity	0,0014*	1,21e-08*	1,55e-05*	< 2,2e-16*	< 2,2e-16*	1,74e-11*
Entropy	0,4226	0,0530	0,2232	< 2,2e-16*	< 2,2e-16*	1,09e-07*
SRE	0,0763	2,75e-06*	6,94e-05*	< 2,2e-16*	< 2,2e-16*	1,30e-10*
LRE	0,0753	2,84e-06*	5,85e-05*	< 2,2e-16*	< 2,2e-16*	1,25e-10*
LGZE	0,0001*	0,4197	1,79e-05*	2,73e-10*	< 2,2e-16*	< 2,2e-16*
HGZE	9,28e-06*	0,1960	0,0276	1,91e-11*	< 2,2e-16*	< 2,2e-16*
SUVmax	4,84e-14*	< 2,2e-16*	< 2,2e-16*	4,84e-14*	< 2,2e-16*	< 2,2e-16*
SUVmean	2,56e-11*	< 2,2e-16*	< 2,2e-16*	2,56e-11*	< 2,2e-16*	< 2,2e-16*
SUVpeak	6,12e-14*	< 2,2e-16*	< 2,2e-16*	6,12e-14*	< 2,2e-16*	< 2,2e-16*
VM	1,0	0,8713	0,8713	1,0	0,8713	0,8713
TLG	0,0001*	4,45e-15*	1,98e-11*	0,0001*	4,45e-15*	1,98e-11*

Tableau 4.2 – P-values du test de Wilcoxon entre le tissu tumoral, le tissu hépatique et le tissu controlatéral dans la cohorte CM avec la discrétisation relative et absolue (entre 0 et 20 unités SUV). Les p-values de VM pour VOI-T vs. VOI-C et VOI-F vs. VOI-C sont inférieures à 1 car la taille de la VOI d'une lésion ne permettait pas son report dans le sein controlatéral. Les p-values en gras sont significatives au seuil de 5%. * signifie que la p-value reste significative après la correction de Holm.

Enfin, nous avons comparé les performances des index de texture calculés avec DA avec celles de ces mêmes index calculés avec DR et combinés avec SUVmax dans une régression logistique. Les résultats sont présentés dans le Tableau 4.5 pour la distinction entre VOI-T et VOI-F sur la cohorte CPNPC-2. On observe que pour 4 index sur 6 la combinaison n'est pas meilleure que l'index seul calculé après une discrétisation absolue. À l'opposé, les p-values sont plus faibles pour LGZE et HGZE calculés avec DR puis combinés avec SUVmax.

Index	VOI-F (CPNPC-2) vs. VOI-F (CM)	VOI-T (CPNPC-2) vs. VOI-T (CM)
Homogeneity	0,6606	0,2643
Entropy	0,1375	0,0005*
SRE	0,0338	0,6365
LRE	0,0075	0,8590
LGZE	0,9439	0,2819
HGZE	0,5485	0,1069
SUVmax	0,1968	0,0237
SUVmean	0,7911	0,1448
SUVpeak	0,0734	0,0118
VM	7,28e-06*	7,28e-06*
TLG	4,45e-06*	2,83e-05*

Tableau 4.3 – P-values du test de Wilcoxon entre la cohorte CPNPC-2 et CM pour le tissu hépatique et tumoral (discrétisation absolue entre 0 et 20 unités SUV). Les p-values en gras sont significatives au seuil de 5%. * signifie que la p-value reste significative après la correction de Holm.

Index	VOI-T vs VOI-F		
	Borne haute = 15 unités SUV	Borne haute = 20 unités SUV	Borne haute = 25 unités SUV
Homogeneity	< 2,2e-16	< 2,2e-16	< 2,2e-16
Entropy	< 2,2e-16	< 2,2e-16	< 2,2e-16
SRE	< 2,2e-16	< 2,2e-16	< 2,2e-16
LRE	< 2,2e-16	< 2,2e-16	< 2,2e-16
LGZE	2,35e-11	1,94e-11	1,53e-09
HGZE	1,92e-12	2,06e-12	2,49e-10

Tableau 4.4 – P-values du test de Wilcoxon entre VOI-T et VOI-F de la cohorte CPNPC-2 suivant la valeur de la borne haute utilisée pour la discrétisation absolue. Les p-values en gras sont significatives au seuil de 5%. * signifie que la p-value reste significative après la correction de Holm.

VOI-T vs. VOI-F		
Index	Index(DA)	Index(DR)+SUVmax
Homogeneity	< 2,2e-16*	2,03e-15*
Entropy	< 2,2e-16*	1,32e-13*
SRE	< 2,2e-16*	5,50e-15*
LRE	< 2,2e-16*	5,50e-15*
LGZE	1,94e-11*	< 2,2e-16*
HGZE	2,06e-12*	3,79e-16*

Tableau 4.5 – P-values du test de Wilcoxon entre VOI-T et VOI-F sur la cohorte CPNPC-2 pour l'index seul calculé après discrétisation absolue (DA) entre 0 et 20 unités SUV et l'index calculé après discrétisation relative (DR) puis combiné avec SUVmax (régression logistique). Les p-values en gras sont inférieures à 5%. * signifie que la p-value reste significative après la correction de Holm.

4.1.3 Discussion

L'utilisation de la discrétisation absolue avant le calcul des matrices de texture permet aux index de différencier les types de tissu sans être affectés par le VM. Nous avons ainsi pu montrer que l'analyse de texture à la tumeur permettait de différencier le tissu tumoral du tissu sain ainsi que le tissu hépatique du tissu controlatéral. Dans les deux groupes de patients étudiés, le tissu controlatéral est plus homogène que le tissu hépatique, qui est lui même plus homogène que le tissu tumoral avec des valeurs d'Homogeneity, LRE et LGZE plus élevées et des valeurs d'Entropy, SRE et HGZE plus faibles. On observe donc des variations intuitives de ces paramètres, qui correspondent bien aux résultats des simulations. Par exemple, l'index Homogeneity augmente entre le tissu tumoral et le tissu sain alors qu'à l'opposé HGZE, qui quantifie les hauts niveaux de gris, diminue. À notre connaissance, seulement deux publications se sont intéressées aux variations des index de texture suivant différents tissus. À partir de 103 patients non traités, avec des lésions osseuses et des tissus mous, [Xu et al., 2014] ont montré que les index de texture issus des images TEP et TDM permettaient de distinguer les tumeurs malignes des lésions bénignes. Par exemple, Entropy mesuré sur les images TEP au 18F-FDG était plus élevé dans les lésions malignes que dans les lésions bénignes. La seconde étude a rapporté que l'index Contrast, issu de la matrice de co-occurrence, était plus élevé dans le foie que dans les tumeurs mammaires de 11 patientes, à partir des images TEP obtenues avec du 18F-FLT avant traitement [Willaime et al., 2013]. Dans notre étude, nous avons observé que ce même index, appartenant au groupe de corrélation n°1, était plus élevé dans les tumeurs mammaires (médiane : 36,4, intervalle : [1,0–251,0]) que dans le tissu hépatique (médiane : 1,0, intervalle : [0,4–2,8]). Ces résultats contradictoires peuvent s'expliquer par la différence entre les traceurs utilisés, le faible nombre de patientes dans [Willaime et al., 2013], l'inclusion de tumeurs avec une large gamme de volumes (de 2,65 à 540,99 ml) sans prendre en compte ces différences volumiques et l'utilisation de la discrétisation relative et non absolue. Il est important de souligner qu'avec la discrétisation relative, aucun index de texture ne conduit à une p-value plus faible que celle de SUVmax, alors qu'avec la méthode absolue, 4 index sur 6 sont meilleurs que le SUVmax pour la distinction VOI-T vs. VOI-F. À l'inverse, toujours dans cette même étude, [Willaime et al., 2013] montrent également que Homogeneity est plus élevé dans le tissu mammaire sain que dans le tissu tumoral ou hépatique, ce qui est bien cohérent avec nos observations.

Dans la comparaison de l'hétérogénéité tumorale, nous avons montré que Entropy était le seul index de texture capable de distinguer les lésions mammaires des lésions pulmonaires avec une p-value égale à 0,0005 et significative même après la correction de Holm. Cet index permet peut-être de capturer une « signature » tumorale caractérisée par une organisation particulière des lésions, ce qui reste à confirmer sur d'autres types tumoraux.

Enfin, même si le choix de la borne supérieure ne semble pas influencer les performances de distinction des index de texture, il peut influencer leur valeur absolue, d'où la nécessité de toujours renseigner avec une très grande précision cette valeur, tout comme l'ensemble des paramètres utilisés pour le calcul des index de texture. Étant donné que les index de texture calculés avec la méthode absolue reflètent une partie de l'information SUV, nous avons comparé les p-values du test de Wilcoxon obtenues pour les index calculés après DA avec celles des index obtenus après DR et combinés avec SUVmax pour la distinction VOI-T vs. VOI-F. Nous avons montré que les index conduisent à des p-values très similaires pour différencier les tissus tumoraux des tissus hépatiques dans la cohorte CPNPC-2. Un avantage d'utiliser la méthode absolue, en plus d'être indépendant du VM, est donc d'inclure à la fois une information sur la texture et la fixation des lésions dans un seul index, ce qui facilite l'utilisation de l'analyse de texture en clinique par rapport à une analyse conjointe de deux paramètres (index de texture après DR combiné à SUVmax).

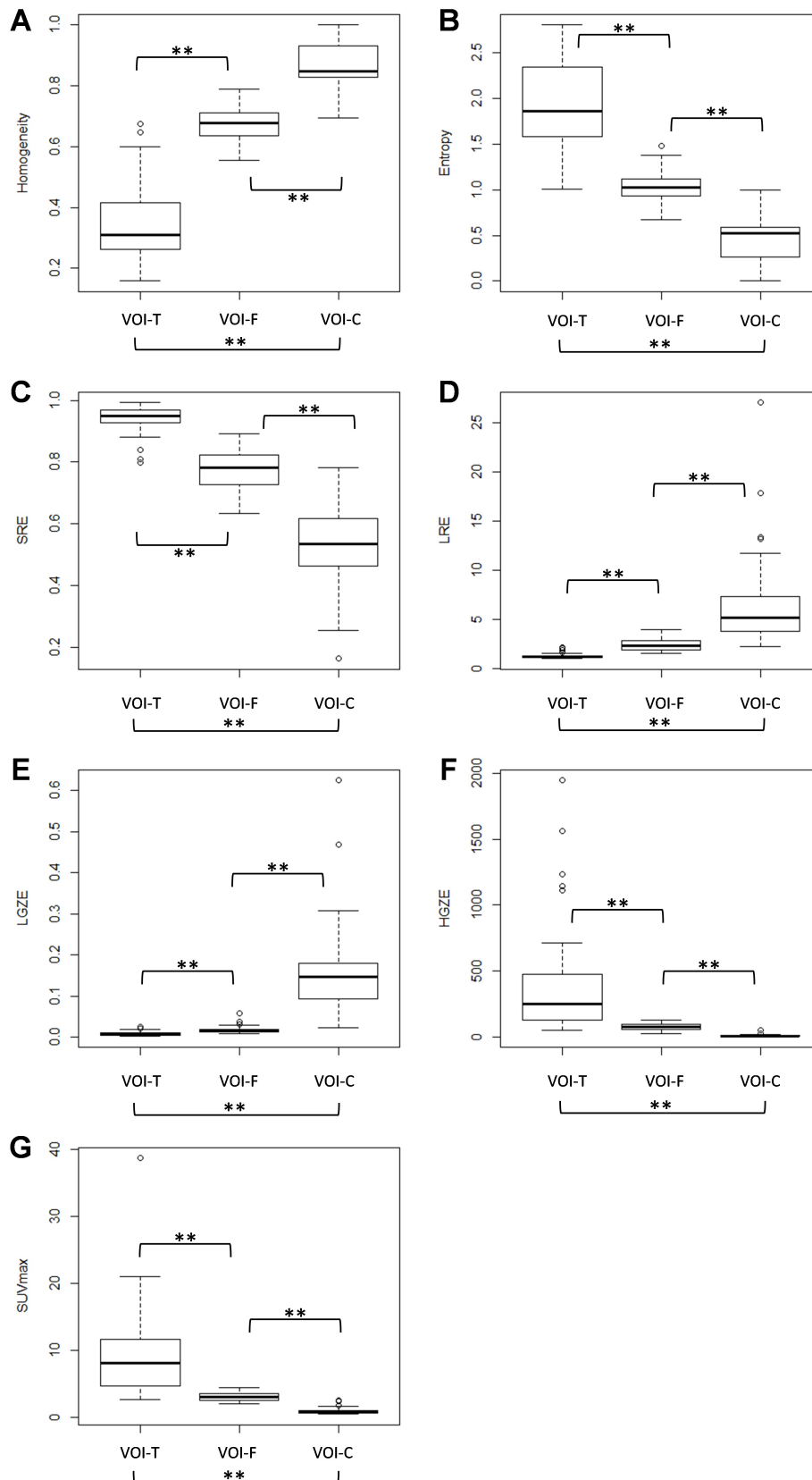


FIGURE 4.1 – Boîtes à moustache pour les index de texture calculés avec DA (entre 0 et 20 unités SUV) et pour SUVmax pour la cohorte CPNPC-2, illustrant la différenciation entre les types de tissus. ** signifie que la différence est significative après la correction de Holm.

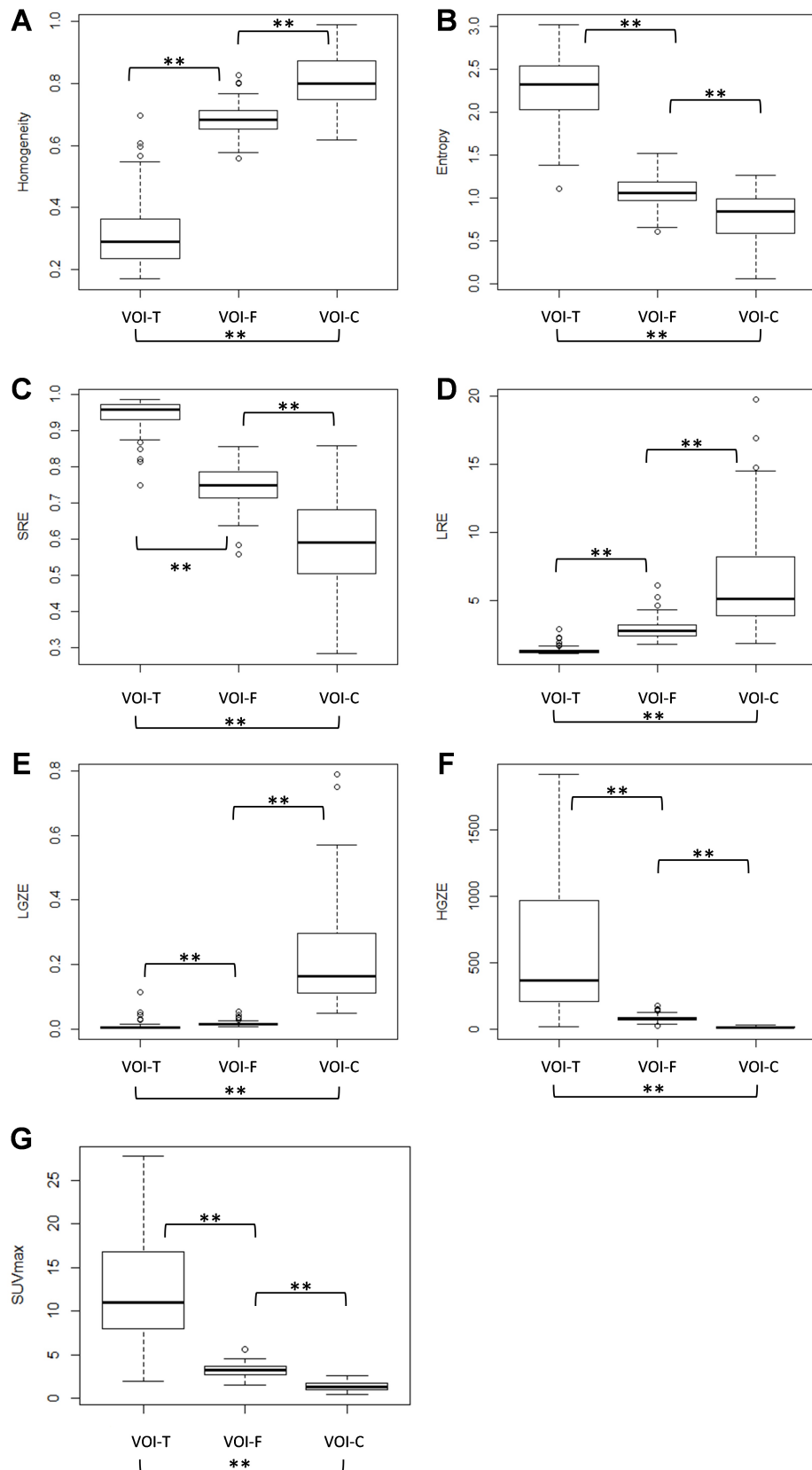


FIGURE 4.2 – Boîtes à moustache pour les index de texture calculés avec DA (entre 0 et 20 unités SUV) et pour SUVmax pour la cohorte CM, illustrant la différenciation entre les types de tissus. ** signifie que la différence est significative après la correction de Holm.

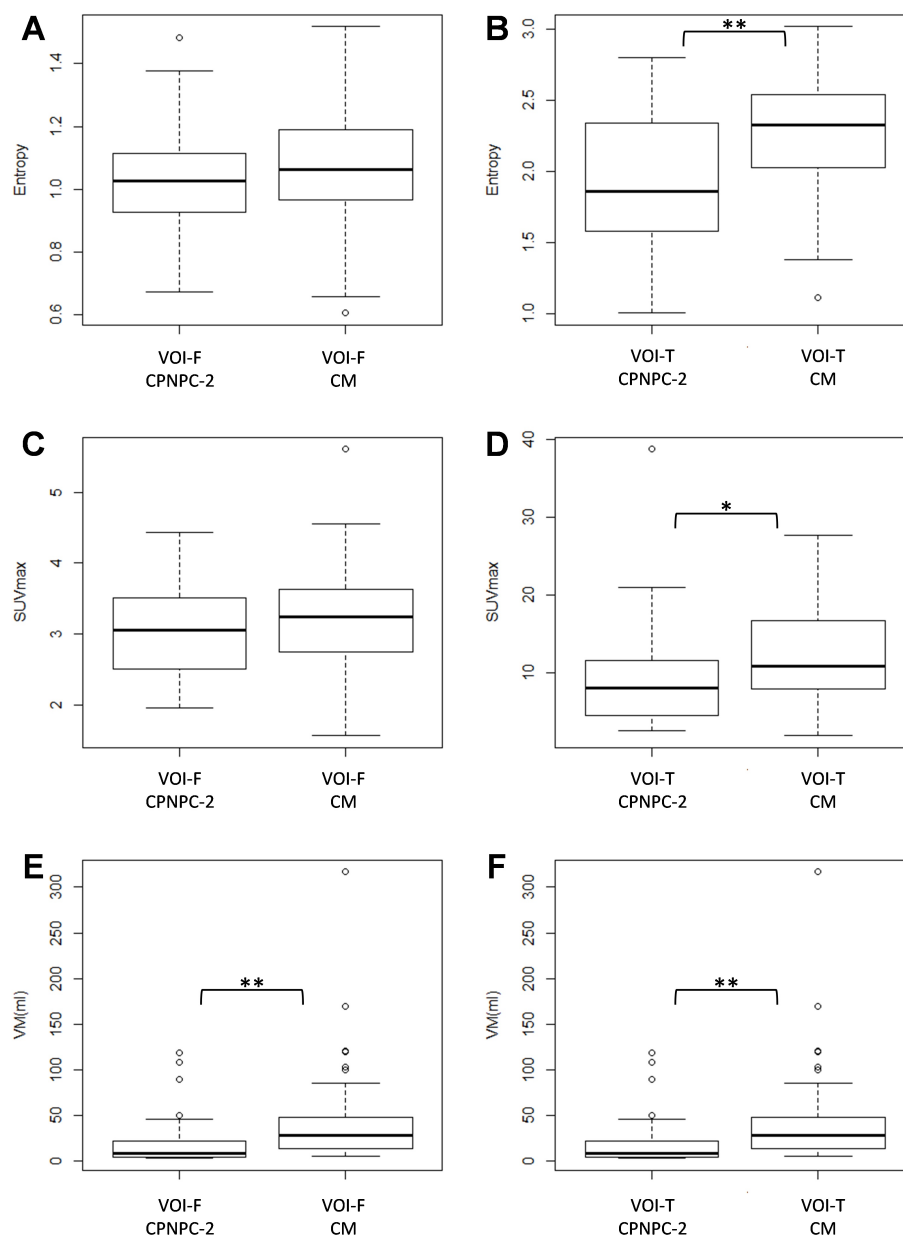


FIGURE 4.3 – Boîtes à moustache pour Entropy (calculé après DA entre 0 et 20 unités SUV), SUVmax et VM entre le tissu hépatique ou le tissu tumoral des deux cohortes. * et ** signifient respectivement que la différence est significative sans ou avec la correction de Holm.

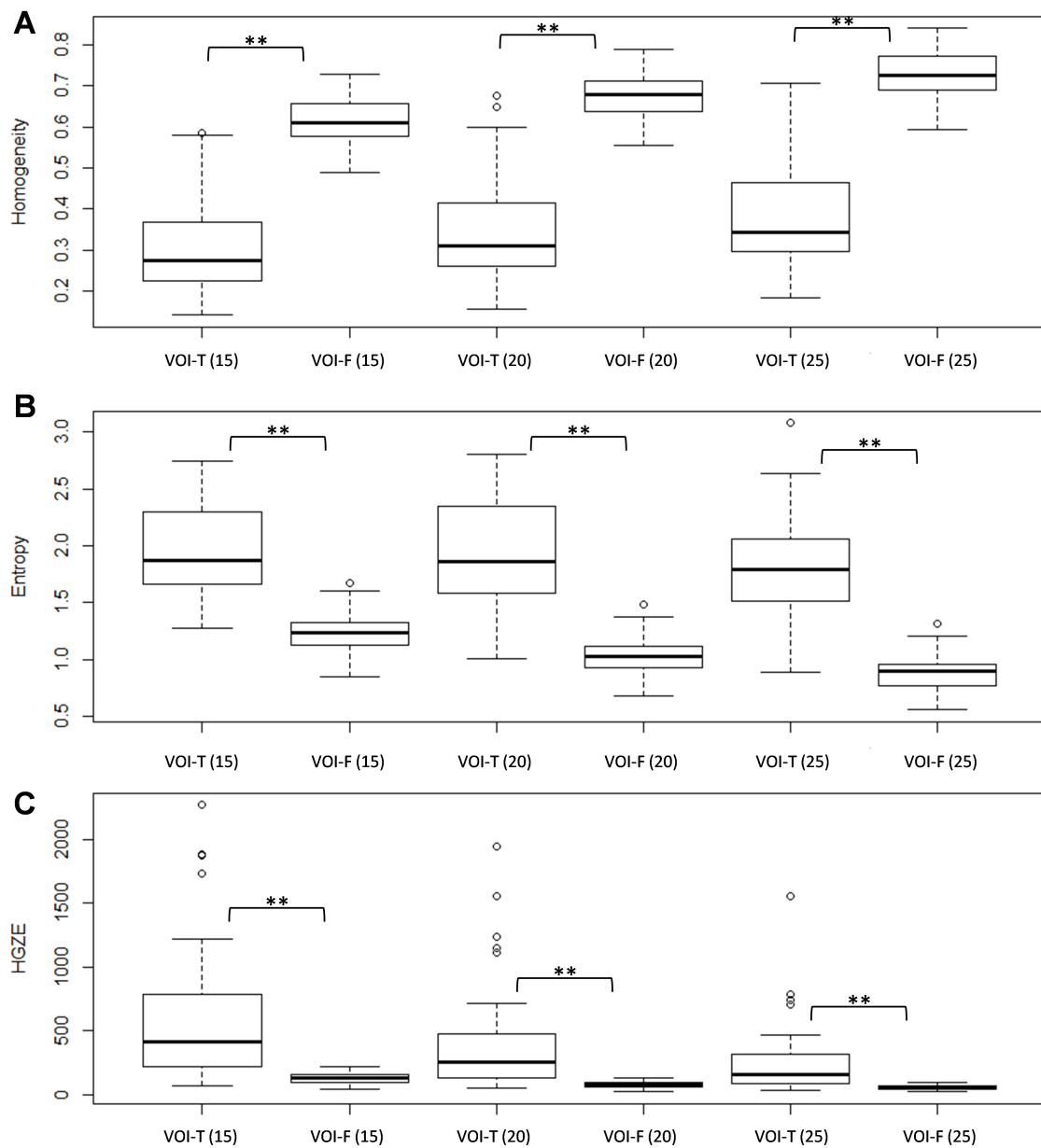


FIGURE 4.4 – Boîtes à moustache pour Homogeneity, Entropy et HGZE entre le tissu hépatique ou le tissu tumoral pour la cohorte CPNPC-2 suivant la valeur de la borne haute utilisée pour la discrétisation absolue (15, 20 ou 25 unités SUV). ** signifie que la différence est significative avec la correction de Holm.

4.2 Corrélation avec l'analyse visuelle

Après avoir démontré que les index de texture calculés avec une discrétisation absolue pouvaient distinguer les différents types de tissus, nous avons étudié la capacité des paramètres à quantifier l'hétérogénéité tumorale telle qu'évaluée visuellement par deux experts.

4.2.1 Matériels & méthodes

Pour comparer l'hétérogénéité tumorale évaluée visuellement et l'hétérogénéité quantifiée par les index de texture, nous avons utilisé les données de la cohorte CM dont les caractéristiques des 54 patientes ont été présentées au Chapitre 3.

Après avoir converti les images TEP au 18F-FDG en unités SUV, l'hétérogénéité de la tumeur primitive a été évaluée visuellement par deux experts indépendamment, en séparant les lésions en deux groupes : les lésions homogènes et hétérogènes. Les experts n'ont pas reçu d'instruction pour classer les lésions dans les deux groupes, le classement a été laissé à leur propre jugement. Les lecteurs n'avaient pas connaissance des informations cliniques des patients et du classement de l'autre lecteur. Les images ont été présentées dans le même ordre aux deux experts et avec la même échelle de couleur (entre 0 et 2 fois SUVmean).

Chaque lésion primitive a ensuite été segmentée avec un seuillage adaptatif à partir des images TEP. Cinq index conventionnels (mesures SUV, VM et TLG) ont été calculés ainsi que six paramètres de texture (Homogeneity, Entropy, SRE, LRE, LGZE et HGZE) après avoir discrétisé les intensités avec une méthode absolue entre 0 et 20 unités SUV et avec 64 niveaux.

Le test du Kappa de Cohen (κ) a été utilisé pour mesurer l'accord entre les deux lecteurs pour la classification des lésions en deux groupes. La capacité de chaque index à séparer les lésions homogènes des lésions hétérogènes, pour chaque lecteur, a été évaluée en utilisant un test de Wilcoxon avec une correction de Holm.

4.2.2 Résultats

La concordance entre les deux lecteurs mesurée par κ est égale à 0,845 ($p\text{-value}=3,26 \times 10^{-10}$) pour séparer les lésions mammaires en lésions homogènes ($n=23$ pour le lecteur n°1 vs. $n=19$ pour le lecteur n°2) et hétérogènes ($n=31$ et $n=35$ respectivement). Les deux lecteurs sont en désaccord sur seulement 4 tumeurs sur 54.

Le Tableau 4.6 présente les p-values du test de Wilcoxon avec la correction de Holm suivant le classement des deux lecteurs séparément ainsi que les résultats pour les 50 tumeurs où le jugement est concordant entre les deux experts. Les p-values obtenues pour les index de texture et les index conventionnels sont proches entre les deux lecteurs. Quand on ne considère que les lésions pour lesquelles il y a une concordance des jugements, les index Entropy, LGZE et HGZE permettent de différencier les deux groupes de lésions avec des p-values inférieures à 5% même après la correction de Holm. Entropy conduit même à une meilleure discrimination des lésions homogènes et hétérogènes telles que qualifiées par les observateurs que les index conventionnels excepté pour le TLG.

Les sens de variations des paramètres sont intuitifs. D'après les Figures 4.5 et 4.6, les lésions classées homogènes ont des valeurs pour Homogeneity, LRE et LGZE plus élevées que pour les lésions hétérogènes. Le comportement contraire est observé pour Entropy, SRE, HGZE et les mesures

SUV, MV et TLG. Les lésions homogènes ont une valeur médiane pour Homogeneity égale à 0,354, Entropy = 1,86, LGZE = 0,0076 et HGZE = 196,3 contre Homogeneity = 0,262, Entropy = 2,48, LGZE = 0,0031 et HGZE = 490,7 pour les lésions classées hétérogènes. Dix exemples de lésions sont présentés sur la Figure 4.7 avec les valeurs associées des index de texture et des index conventionnels.

Index	Lecteur n°1	Lecteur n°2	Accord entre les 2 lecteurs
Homogeneity	0,0561	0,0103*	0,0173
Entropy	2,68e-08*	1,27e-09*	9,62e-10*
SRE	0,1333	0,0436	0,0571
LRE	0,1333	0,0643	0,0718
LGZE	0,0025*	2,32E-05*	6,44e-05*
HGZE	0,0012*	6,82e-06*	1,82e-05*
SUVmax	0,0004*	7,69e-07*	2,41e-06*
SUVmean	0,0008*	4,94e-06*	1,30e-05*
SUVpeak	0,0005*	9,83e-07*	3,52e-06*
VM	2,70e-07*	3,86e-06*	1,61e-06*
TLG	5,00e-12*	1,09e-14*	6,58e-14*

Tableau 4.6 – P-values du test de Wilcoxon pour la distinction entre les lésions classées homogènes et hétérogènes pour chacun des deux lecteurs (discretisation absolue entre 0 et 20 unités SUV). La dernière colonne correspond aux p-values du test de Wilcoxon pour les lésions où les deux lecteurs sont d'accord sur la nature de la tumeur (homogène ou hétérogène). Les p-values en gras sont significatives au seuil de 5%. * signifie que la p-value reste significative après la correction de Holm.

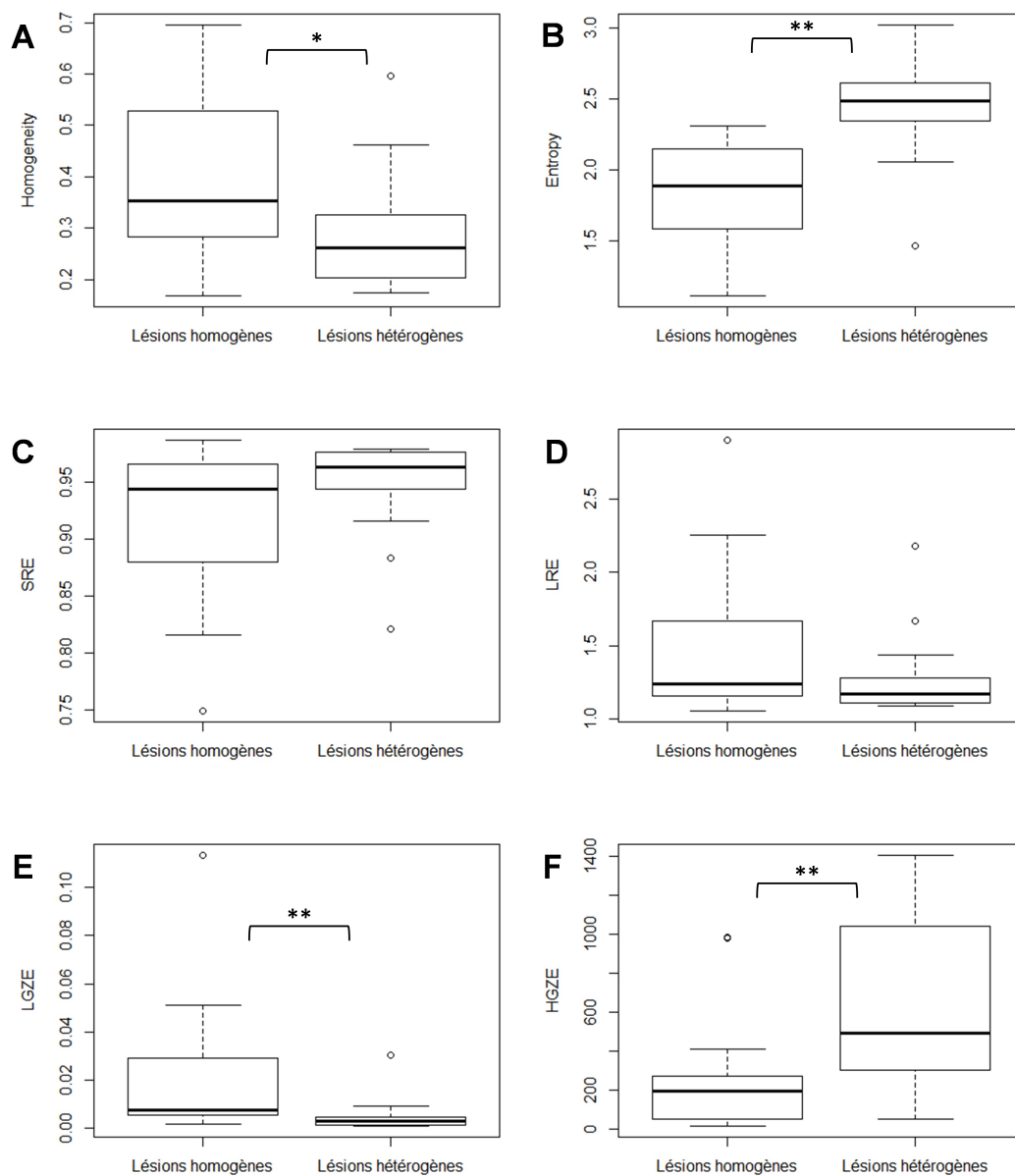


FIGURE 4.5 – Boîtes à moustache pour les index de texture entre les lésions classées homogènes et hétérogènes par les deux lecteurs simultanément (discrétisation absolue entre 0 et 20 unités SUV). * et ** signifient respectivement que la différence est significative sans ou avec la correction de Holm.

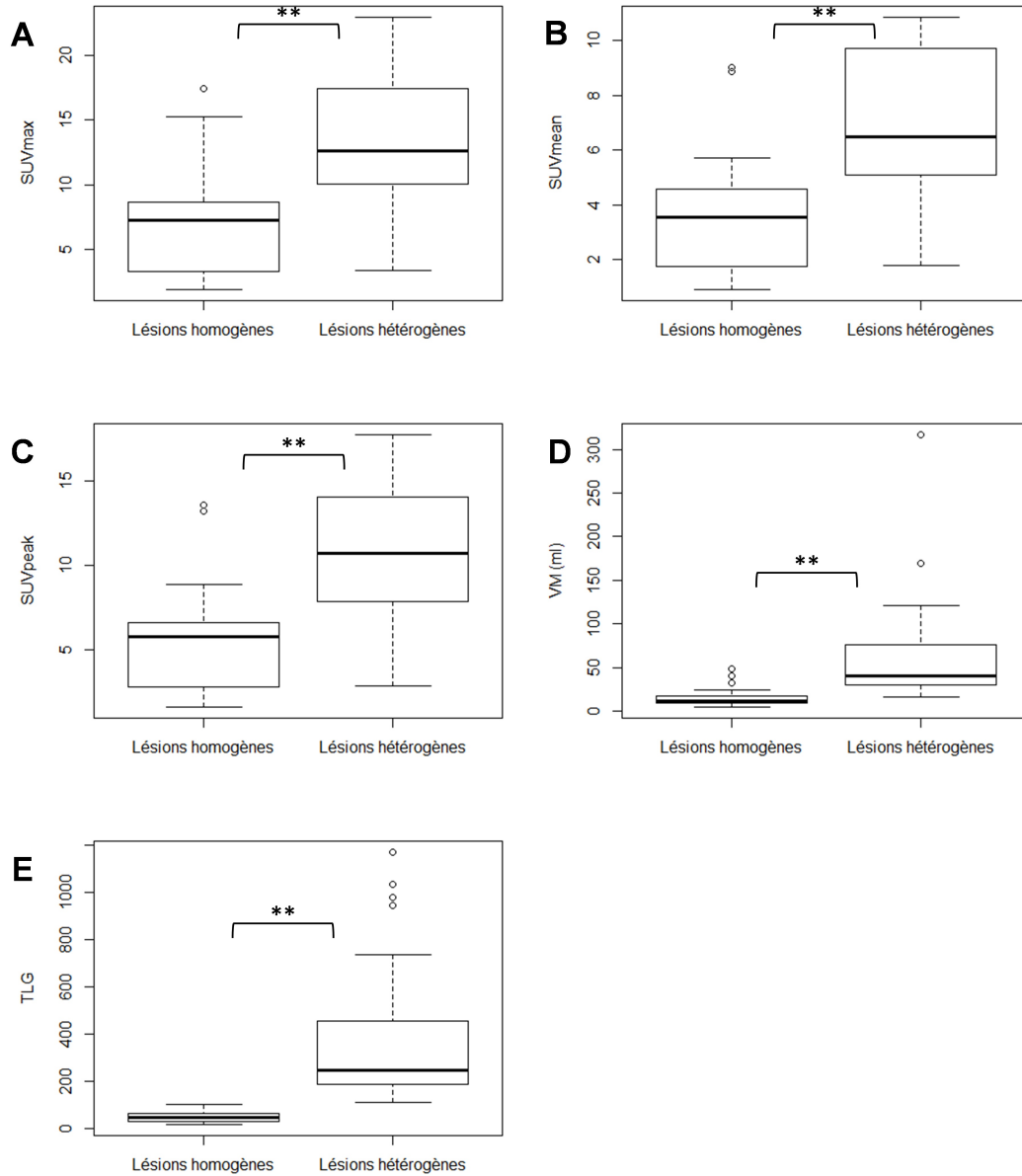


FIGURE 4.6 – Boîtes à moustache pour les index conventionnels entre les lésions classées homogènes et hétérogènes par les deux lecteurs simultanément. * et ** signifient respectivement que la différence est significative sans ou avec la correction de Holm.

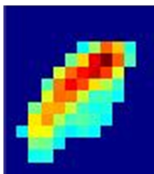
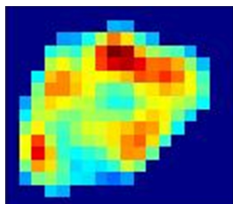
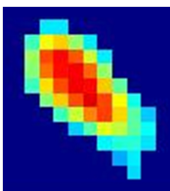
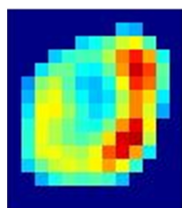
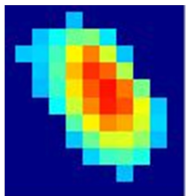
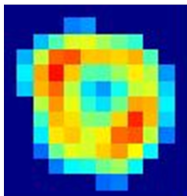
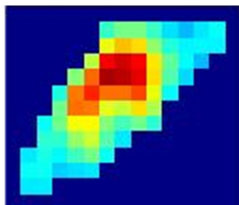
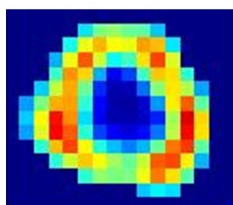
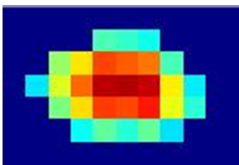
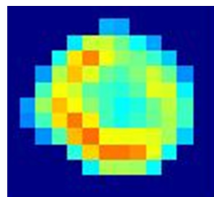
Lésions classées homogènes par les deux lecteurs		Lésions classées hétérogènes par les deux lecteurs	
	Homogeneity : 0,607 Entropy : 1,38 SRE : 0,816 LRE : 2,25 LGZE : 0,0509 HGZE : 32,7 SUVmax : 3,0 SUVmean : 1,4 SUVpeak : 2,3 VM : 48,3 ml TLG : 70,0		Homogeneity : 0,367 Entropy : 2,26 SRE : 0,932 LRE : 1,33 LGZE : 0,0090 HGZE : 177,4 SUVmax : 8,0 SUVmean : 3,7 SUVpeak : 5,9 VM : 66,5 ml TLG : 243,8
	Homogeneity : 0,520 Entropy : 1,65 SRE : 0,874 LRE : 1,69 LGZE : 0,0295 HGZE : 51,7 SUVmax : 3,6 SUVmean : 1,9 SUVpeak : 3,1 VM : 32,3 ml TLG : 60,0		Homogeneity : 0,309 Entropy : 2,47 SRE : 0,949 LRE : 1,23 LGZE : 0,0040 HGZE : 364,2 SUVmax : 10,9 SUVmean : 5,5 SUVpeak : 8,4 VM : 57,3 ml TLG : 316,2
	Homogeneity : 0,565 Entropy : 1,41 SRE : 0,851 LRE : 1,88 LGZE : 0,0284 HGZE : 48,4 SUVmax : 2,6 SUVmean : 1,4 SUVpeak : 2,1 VM : 20,3 ml TLG : 28,7		Homogeneity : 0,189 Entropy : 2,68 SRE : 0,978 LRE : 1,09 LGZE : 0,0014 HGZE : 1103,3 SUVmax : 17,0 SUVmean : 9,8 SUVpeak : 13,8 VM : 30,7 ml TLG : 301,9
	Homogeneity : 0,547 Entropy : 1,52 SRE : 0,867 LRE : 1,72 LGZE : 0,0314 HGZE : 50,7 SUVmax : 3,1 SUVmean : 1,6 SUVpeak : 2,5 VM : 14,1 ml TLG : 23,3		Homogeneity : 0,174 Entropy : 3,02 SRE : 0,978 LRE : 1,10 LGZE : 0,0018 HGZE : 1405,6 SUVmax : 22,9 SUVmean : 10,8 SUVpeak : 17,7 VM : 67,9 ml TLG : 735,2
	Homogeneity : 0,536 Entropy : 1,47 SRE : 0,886 LRE : 1,64 LGZE : 0,0447 HGZE : 35,8 SUVmax : 3,0 SUVmean : 1,5 SUVpeak : 2,2 VM : 10,8 ml TLG : 16,1		Homogeneity : 0,203 Entropy : 2,77 SRE : 0,976 LRE : 1,11 LGZE : 0,0014 HGZE : 1079,2 SUVmax : 18,2 SUVmean : 9,8 SUVpeak : 14,7 VM : 43,6 ml TLG : 429,3

FIGURE 4.7 – Exemple de dix lésions classées soit homogènes, soit hétérogènes par les deux lecteurs avec les index de texture (calculés après DA entre 0 et 20 unités SUV) et les index conventionnels associés.

4.2.3 Discussion

Afin de valider le sens des variations des index de texture, deux experts ont évalué visuellement l'hétérogénéité tumorale de 54 lésions mammaires à partir des images TEP au 18F-FDG. L'accord entre les deux lecteurs est très bon ; toutefois dans 7% des cas, on observe un désaccord entre leur jugement.

Nos résultats montrent que certains index de texture permettent de retrouver la classification des experts et qu'ils varient intuitivement : les tumeurs hétérogènes ont des valeurs d'Entropy et de HGZE plus élevées que pour les tumeurs homogènes ; le constat inverse peut être fait pour Homogeneity et LGZE. Ces variations obtenues après une discrétisation absolue des niveaux de gris dans l'image sont contraires à celles rapportées par [Tixier et al., 2014b]. Les auteurs montrent que l'index Homogeneity, calculé après une discrétisation relative, varie contre-intuitivement avec la classification de deux experts (cf. Figure 3.13), démontrant à nouveau l'intérêt d'utiliser la discrétisation absolue.

En plus des index conventionnels, quatre index permettent de retrouver la classification des lecteurs (avec en plus Homogeneity pour le classement du lecteur n°2). L'index Entropy conduit aux meilleures performances par rapport aux autres index de texture, aux mesures SUV et à VM. Cependant, la mesure du TLG conduit à une p-value plus faible que celle de Entropy.

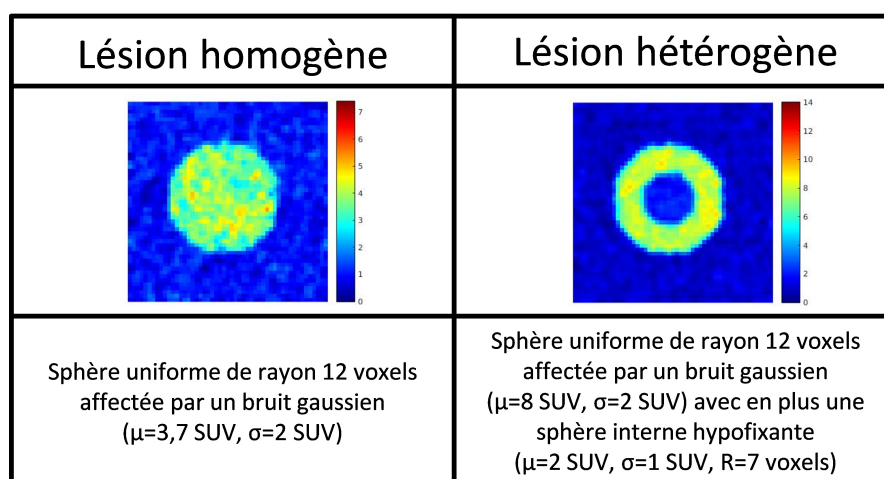


FIGURE 4.8 – Caractéristiques des simulations de lésions homogènes et hétérogènes.

Pour valider les simulations décrites au Chapitre 3, nous avons construit deux nouveaux modèles d'après les caractéristiques des tumeurs mammaires étudiées (cf. Figure 4.8). Le modèle de lésion homogène est similaire à celui de la sphère #1 simulée à la section 3.5, avec un SUVmean correspondant à celui observé pour les tumeurs mammaires classées homogènes par les deux experts simultanément soit $SUV_{mean} = 3,7$. Le modèle de lésion hétérogène reprend les caractéristiques de la sphère #12 avec un SUVmean égal à 7,0 unités SUV, qui correspond au SUVmean moyen parmi les lésions classées hétérogènes par les deux observateurs. Les boîtes à moustache présentées sur la Figure 4.9 montrent bien que l'on observe le même type de variations pour les simulations que pour les tumeurs mammaires : Entropy, SRE, HGZE et les mesures SUV sont plus élevés pour les tumeurs hétérogènes alors que Homogeneity, LRE et LGZE sont plus élevés pour les tumeurs homogènes. Les simulations apparaissent toutefois plus homogènes que les tumeurs réelles avec, par exemple, Entropy égal à 2,06 (en médiane) pour les simulations hétérogènes et à 1,65 pour les simulations homogènes contre 2,48

pour les lésions hétérogènes et 1,89 pour les lésions homogènes. Pour LGZE et HGZE, leurs valeurs sont similaires entre les simulations et les données réelles, ce qui provient sans doute du fait que les simulations ont été construites avec une fixation moyenne (SUV_{mean}) correspondant à celle observée pour les tumeurs mammaires.

Cette étude nous a permis de montrer que les index de texture variaient intuitivement et pouvaient quantifier l'hétérogénéité de la fixation du traceur observée sur les images TEP.

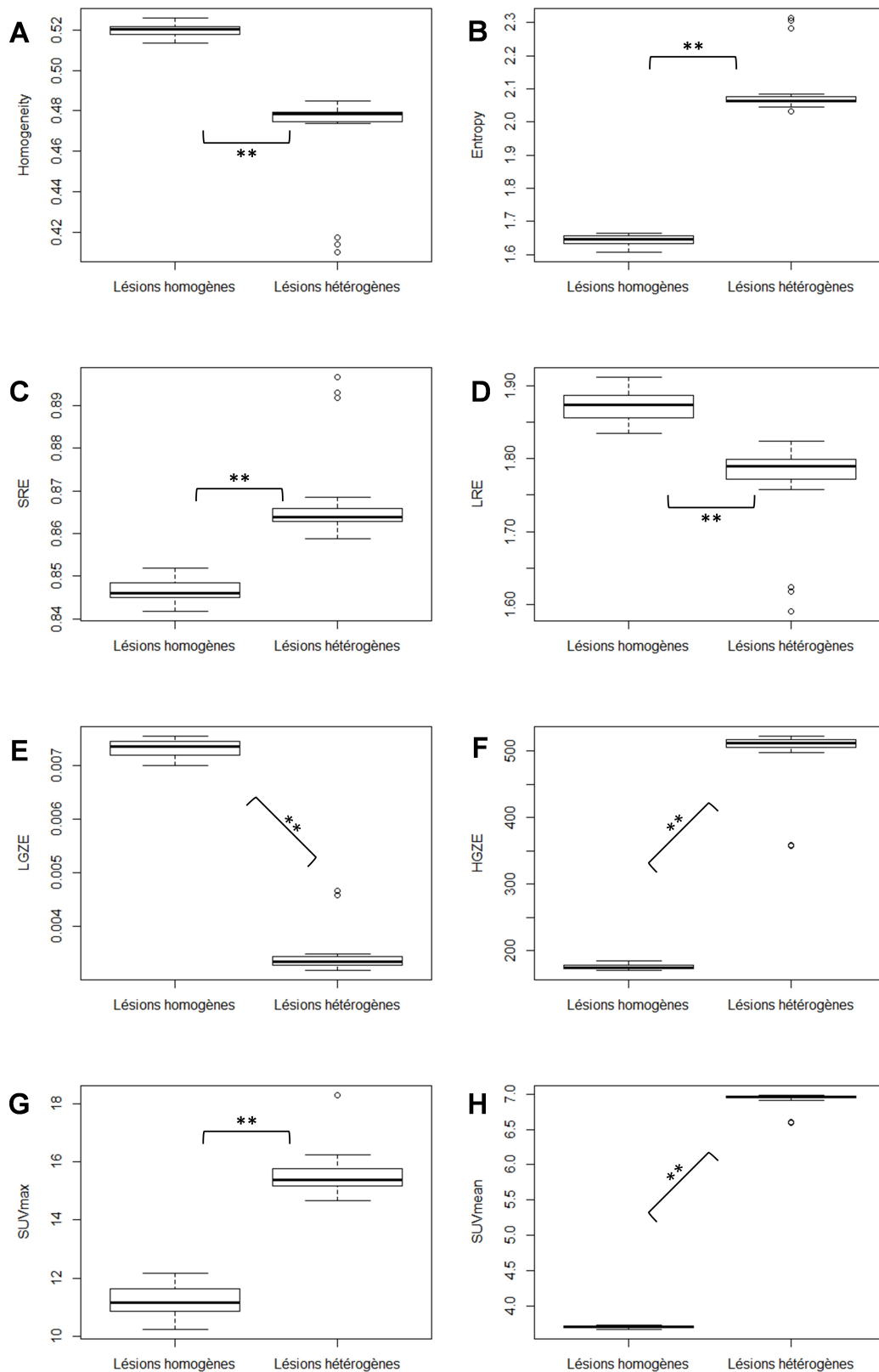


FIGURE 4.9 – Boîtes à moustache pour les index de texture (calculés après DA entre 0 et 20 unités SUV) et pour SUVmax et SUVmean entre les simulations de lésions homogènes et hétérogènes. * et ** signifient respectivement que la différence est significative sans ou avec la correction de Holm.

4.3 Corrélation avec les caractéristiques tumorales

Après avoir étudié la capacité des index de texture à distinguer les tissus sains des tissus tumoraux et avoir montré que ces paramètres étaient bien liés à l'analyse visuelle de l'hétérogénéité tumorale, nous avons testé leur corrélation avec les caractéristiques tumorales pour des lésions mammaires, des lésions pulmonaires et des gliomes.

4.3.1 Les lésions mammaires

Les cancers mammaires sont répartis en quatre classes suivant les caractéristiques histologiques des lésions [Vuong et al., 2014] :

- Luminal A, qui se caractérise par la présence de récepteurs aux œstrogènes et/ou à la progestérone, l'absence de sur-expression de la protéine HER-2 (récepteur transmembranaire impliqué dans la régulation de la prolifération cellulaire) et une prolifération faible ;
- Luminal B, qui est positif pour les récepteurs aux œstrogènes et/ou à la progestérone avec le plus souvent une sur-expression de HER-2 et une prolifération élevée ;
- HER-2 qui présente une sur-expression de HER-2 ;
- Triple-négatif qui est négatif pour les récepteurs aux œstrogènes et à la progestérone et sans sur-expression de HER-2.

L'appartenance à une de ces classes influence la prise en charge du patient dans la stratégie thérapeutique à adopter. Nous avons donc étudié si les index de texture issus des images TEP étaient sensibles aux caractéristiques tumorales permettant l'identification de ces sous-types tumoraux parmi les tumeurs mammaires.

4.3.1.1 Matériels & méthodes

Pour ce faire, nous avons utilisé les données de la cohorte CM dont les caractéristiques ont été détaillées au Chapitre 3. Après un seuillage adaptatif des lésions sur les images TEP au 18F-FDG, nous avons calculé les cinq index conventionnels ainsi que six paramètres de texture (discrétisation absolue entre 0 et 20 unités SUV avec 64 niveaux).

Une biopsie a été pratiquée sur l'ensemble des 54 patientes avant chimiothérapie ou chirurgie. Une analyse histologique a ensuite été réalisée afin de déterminer la présence de récepteurs aux œstrogènes et à la progestérone ainsi que la sur-expression de la protéine HER-2. À l'aide d'un test de Wilcoxon avec la correction de Holm, nous avons étudié la capacité des index de texture à identifier ces caractéristiques tumorales ainsi que les lésions triple-négatives des non triples-négatives. Les effectifs de chaque groupe sont donnés dans le Tableau 4.7.

4.3.1.2 Résultats

Les capacités des index de texture à identifier les caractéristiques des lésions mammaires ont été évaluées grâce à un test de Wilcoxon, dont les p-values sont rapportées dans le Tableau 4.8. LGZE et HGZE, en plus des mesures SUV, permettent de distinguer la présence de récepteurs aux œstrogènes, même après la correction de Holm des p-values. Par exemple, HGZE et SUVmax conduisent à des valeurs plus élevées dans les lésions négatives pour les récepteurs aux œstrogènes, à l'inverse de LGZE (cf. Figure 4.10). Pour la présence de récepteurs à la progestérone, seuls Entropy et TLG sont capables de faire cette distinction, mais seulement avant d'avoir corrigé les p-values des comparaisons multiples. Enfin aucun index ne permet d'identifier les lésions sur-exprimant la protéine HER-2.

Caractéristiques	n
Récepteurs aux œstrogènes	
positif	38
négatif	16
Récepteurs à la progestérone	
positif	19
négatif	35
Sur-expression de HER-2	
positif	13
négatif	41
Types de lésions	
Lésion triple-négative	13
Lésion non triple-négative	41

Tableau 4.7 – Caractéristiques des lésions mammaires.

Index	Récepteurs aux œstrogènes	Récepteurs à la progestérone	Surexpression de HER-2	Tumeurs triple-négatives
Homogeneity	0,0389	0,5062	0,4114	0,0593
Entropy	0,0447	0,0364	0,2278	0,0198
SRE	0,0492	0,6151	0,3670	0,1017
LRE	0,0702	0,6151	0,4000	0,1153
LGZE	0,0023*	0,1434	0,7643	0,0014*
HGZE	0,0023*	0,1385	0,6893	0,0013*
SUVmax	0,0030*	0,1337	0,6600	0,0024*
SUVmean	0,0023*	0,0853	0,6456	0,0009*
SUVpeak	0,0047*	0,1115	0,5891	0,0021*
VM	0,9321	0,1831	0,1567	0,5305
TLG	0,1282	0,0103	0,0715	0,0248

Tableau 4.8 – P-values du test de Wilcoxon pour caractériser la capacité des index à identifier la présence de récepteurs (œstrogène ou progestérone), la sur-expression de HER-2, et les tumeurs triple-négatives (index de texture calculés après DA entre 0 et 20 unités SUV). Les p-values en gras sont significatives au seuil de 5%. * signifie que la p-value reste significative après la correction de Holm.

LGZE, HGZE et les mesures SUV (cf. Tableau 4.8) permettent de distinguer les lésions triple-négatives des lésions non triple-négatives. Les lésions triple-négatives apparaissent plus hétérogènes et plus fixantes que les lésions non triple-négatives (cf. Figure 4.12) avec HGZE et SUVmax plus élevés et LGZE plus faible dans les lésions triple-négatives (cf. Figure 4.11).

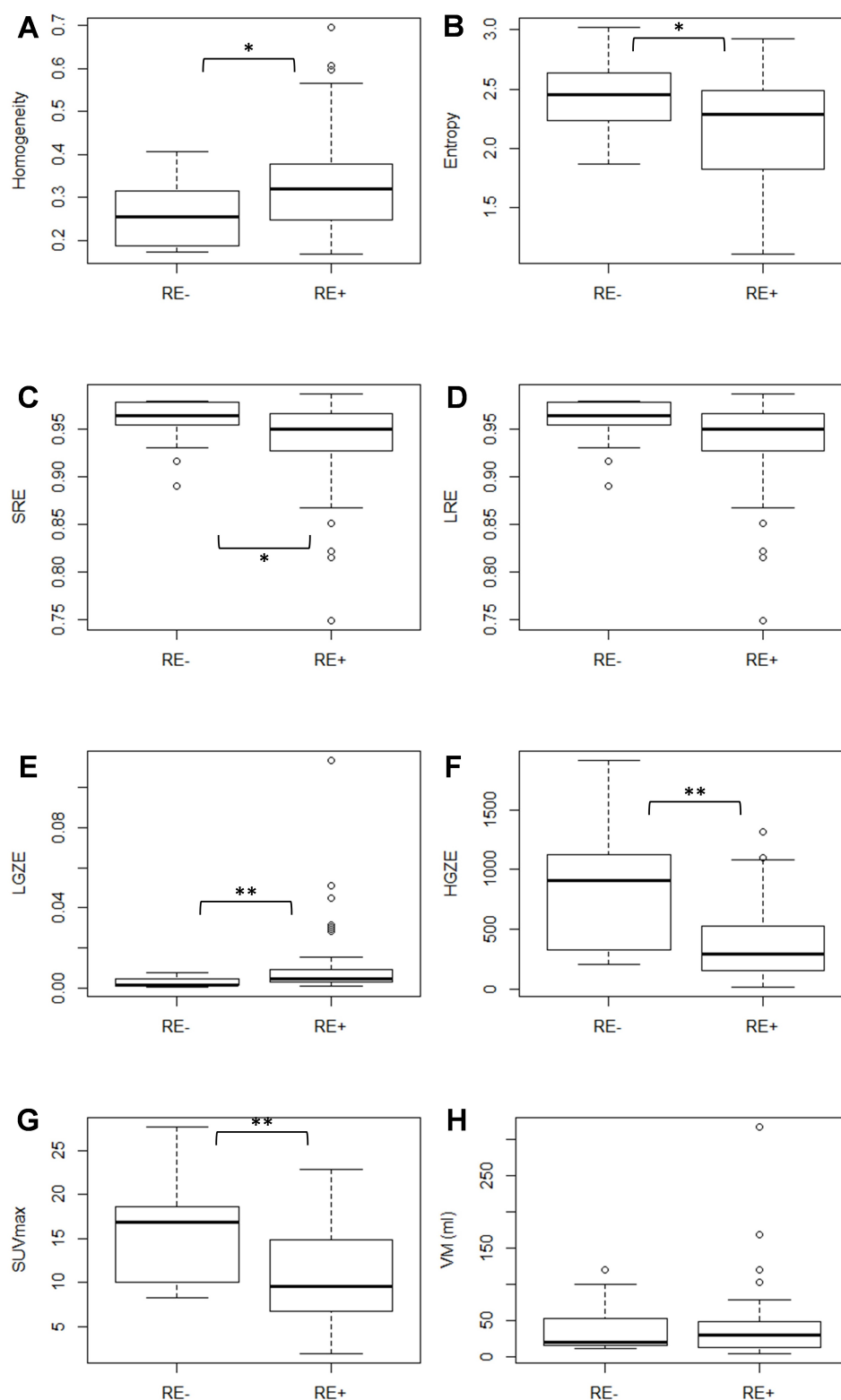


FIGURE 4.10 – Boîtes à moustache entre les lésions positives aux œstrogènes (RE+) et les lésions négatives aux œstrogènes (RE-) pour les index de texture (calculés après DA entre 0 et 20 unités SUV), pour SUVmax et VM. * et ** signifient respectivement que la différence est significative sans ou avec la correction de Holm.

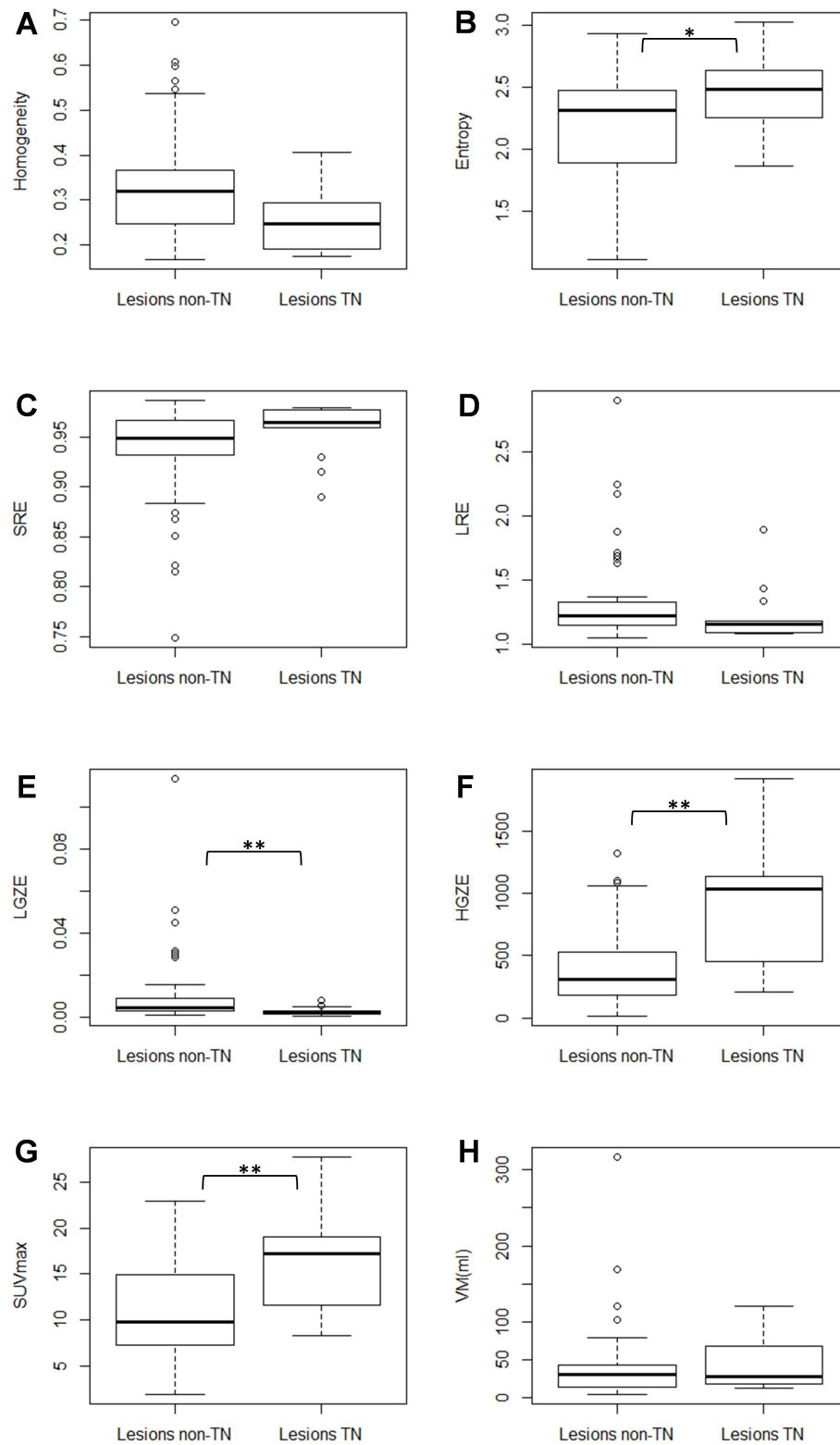


FIGURE 4.11 – Boîtes à moustache entre les lésions non triple-négatives (non-TN) et triple-négatives (TN) pour les index de texture (calculés après DA entre 0 et 20 unités SUV), pour SUVmax et VM. * et ** signifient respectivement que la différence est significative sans ou avec la correction de Holm.

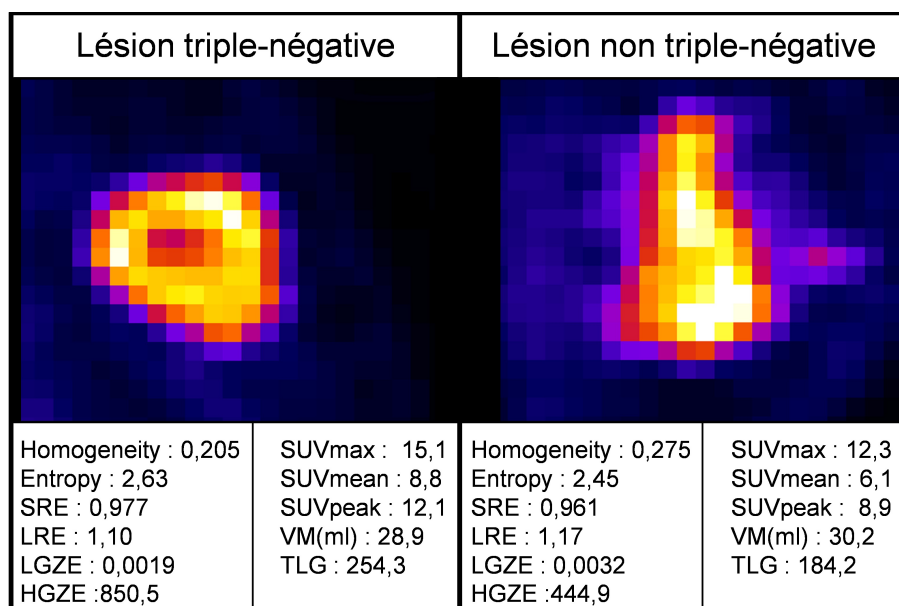


FIGURE 4.12 – Exemple d’une lésion triple-négative et d’une lésion non triple-négative avec les index de texture associés (calculés après DA entre 0 et 20 unités SUV).

4.3.1.3 Discussion

L’identification des caractéristiques tumorales oriente la prise en charge du patient et permet de proposer le traitement le plus approprié. Nous avons montré ici que les paramètres de texture étaient sensibles à la présence de récepteurs aux œstrogènes. Toutefois, l’analyse de texture des images TEP ne semble pas être capable de capturer une « signature » biologique de la sur-expression de HER-2 dans les lésions.

Les tumeurs mammaires triple-négatives sont des tumeurs agressives et de mauvais pronostic [Fornier and Fumoleau, 2012, Metzger-Filho et al., 2012]. [Ahmed et al., 2013] ont montré, sur des images IRM après injection de produit de contraste, une texture différente entre les lésions triple-négatives et les autres types de tumeurs. Également à partir des données IRM, [Uematsu, 2011] a mis en évidence que les lésions triple-négatives sont associées avec la présence d’une nécrose tumorale. Ces résultats sont cohérents avec nos observations puisque nous avons montré que les lésions triple-négatives étaient plus hétérogènes que les autres lésions. De plus, les index LGZE et HGZE conduisent à des p-values plus faibles que celle de Entropy pour la distinction des types tumoraux. Dans le Chapitre 3, à partir des simulations, nous avons montré que LGZE et HGZE étaient plus sensibles aux macro-hétérogénéités, comme les nécroses par exemple. Pour l’identification des lésions triple-négatives, qui auraient tendance à davantage nécroser que les autres lésions, il est donc cohérent que ces deux index soient plus performants que Entropy qui lui est plutôt sensible aux micro-hétérogénéités.

4.3.2 Les lésions pulmonaires non à petites cellules

Nous avons montré que certains index de texture étaient sensibles à certaines caractéristiques tumorales pour les lésions mammaires. Nous avons également évalué la capacité des index de texture à distinguer les caractéristiques d'un second type de cancer, le cancer pulmonaire non à petites cellules.

4.3.2.1 Matériels & méthodes

Afin d'étudier la corrélation entre les index de texture issus des images et les caractéristiques des lésions, nous avons utilisé les données acquises pour la cohorte CPNPC-2, qui a été présentée au Chapitre 3. Après un seuillage adaptatif des lésions sur les images TEP au 18F-FDG avant chirurgie, nous avons calculé les cinq index conventionnels ainsi que six paramètres de texture (discrétisation absolue entre 0 et 20 unités SUV avec 64 niveaux).

Les 48 patients de la cohorte CPNPC-2 ont été opérés et une analyse histologique de la pièce opératoire a été effectuée afin de déterminer le sous-type de cancer, le niveau de différenciation des cellules tumorales et la présence ou non d'une nécrose histologique (cf. Tableau 4.9).

Caractéristiques	n
Stade tumoral	
T1	12
T2	28
autres	8
Stade ganglionnaire	
N0	33
N+	15
Sous-type de cancer	
adénocarcinome	28
épidermoïde	13
autres	7
Niveau de différenciation	
bien différencié	23
peu ou moyennement différencié	11
sans information	14
Présence de nécrose histologique	
oui	31
non	14
sans information	3

Tableau 4.9 – Caractéristiques des lésions pulmonaires non à petites cellules.

Un test de Wilcoxon a été réalisé avec une correction de Holm pour étudier si les index étaient capables d'évaluer le stade tumoral, le stade ganglionnaire ainsi que les critères histologiques identifiés sur la pièce opératoire.

4.3.2.2 Résultats

Le Tableau 4.10 rapporte les p-values du test de Wilcoxon pour les index en fonction des caractéristiques des tumeurs pulmonaires. VM et TLG sont les seuls index capables de distinguer les tumeurs de stade T1 et T2 ainsi que les patients avec une atteinte ganglionnaire.

Seuls Homogeneity et Entropy conduisent à des p-values significatives pour le test de Wilcoxon avec la correction de Holm entre les adénocarcinomes et les épidermoïdes, les deux sous-types les plus présents dans notre cohorte. Les adénocarcinomes présentent une texture plus homogène que les épidermoïdes avec des valeurs d'Homogeneity plus élevées et d'Entropy plus faibles (cf. Figures 4.16 et 4.13).

Concernant le niveau de différenciation, nous avons étudié si les index étaient capables de séparer les lésions bien différenciées des lésions peu ou moyennement différenciées. Entropy est capable de faire cette distinction, tout comme SUVmax et TLG. Entropy conduit même à la plus faible p-value. Les lésions peu ou moyennement différenciées présentent une texture plus hétérogène que les tumeurs bien différenciées (cf. Figure 4.14).

L'analyse de la pièce opératoire a permis de déterminer si les lésions étaient nécrosées ou non. Nous avons montré que tous les index étaient sensibles à la présence de nécrose (cf. Tableau 4.10). Comme l'illustre la Figure 4.15, les lésions non-nécrosées présentent une texture plus homogène que les lésions nécrosées.

Index	Stade tumoral	Stade ganglionnaire	Sous-type	Niveau de différenciation	Nécrose
Homogeneity	0,4055	0,9825	0,0039*	0,0651	5,09e-05*
Entropy	0,0077	0,0071	0,0039*	0,0012*	2,06e-07*
SRE	0,6943	0,6599	0,0179	0,3256	0,0015*
LRE	0,9420	0,3206	0,0506	0,5372	0,0085*
LGZE	0,1298	0,1562	0,0083	0,0230	4,12e-06*
HGZE	0,1298	0,1697	0,0076	0,0187	2,95e-06*
SUVmax	0,1224	0,1150	0,0108	0,0051*	1,22e-06*
SUVmean	0,1376	0,1150	0,0069	0,0150	2,48e-06*
SUVpeak	0,0646	0,0523	0,0076	0,0095	4,85e-06*
VM	0,0009*	0,0018*	0,1303	0,0111	0,0078*
TLG	0,0002*	0,0017*	0,0152	0,0026*	1,07e-05*

Tableau 4.10 – P-values du test de Wilcoxon pour caractériser la capacité des index à identifier les caractéristiques histologiques des lésions pulmonaires (index de texture calculés après DA entre 0 et 20 unités SUV). Les p-values en gras sont significatives au seuil de 5%. * signifie que la p-value reste significative après la correction de Holm.

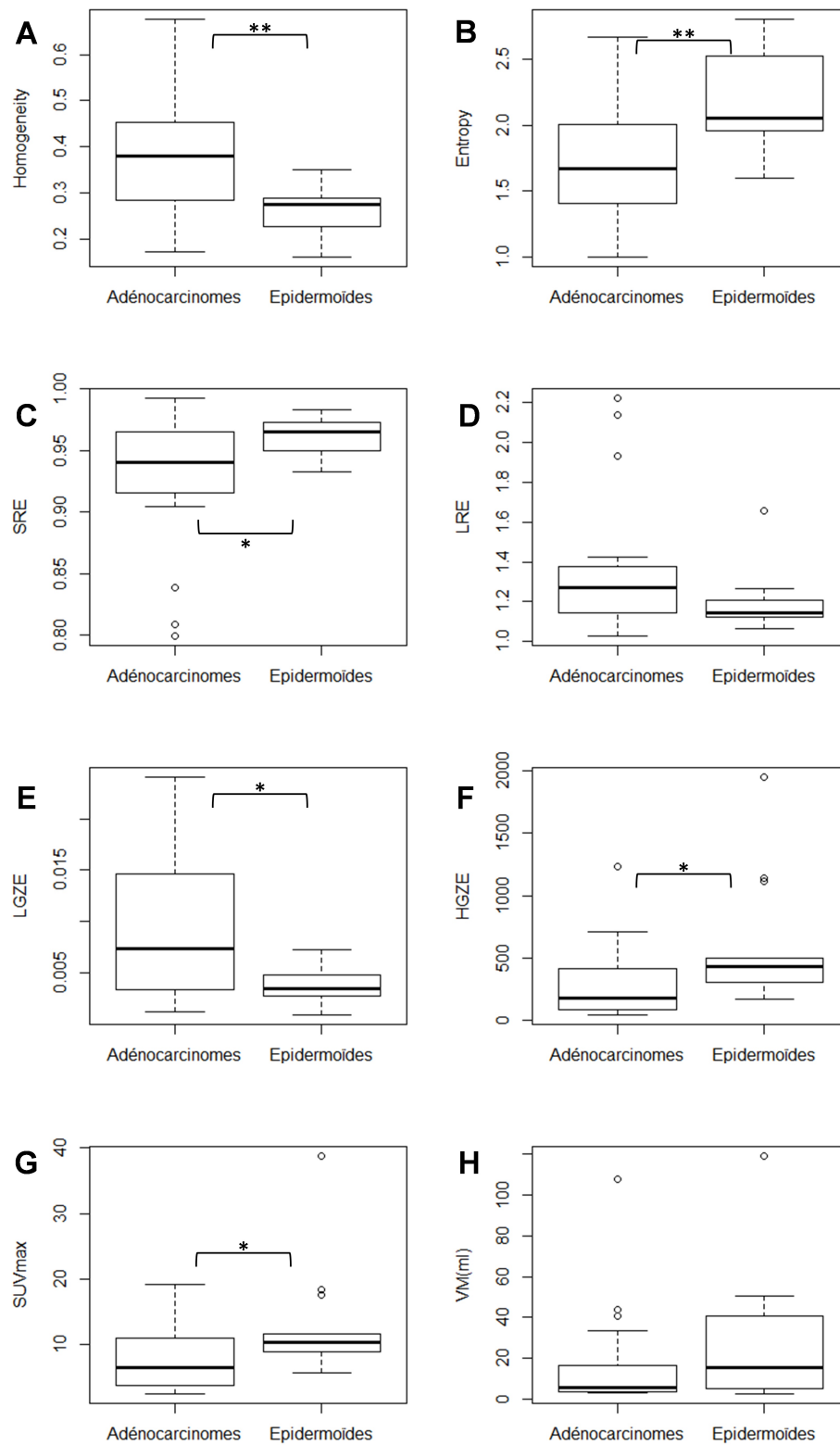


FIGURE 4.13 – Boîtes à moustache entre les lésions de type adénocarcinome et les lésions de type épidermoïde (index de texture calculés entre 0 et 20 unités SUV). * et ** signifient respectivement que la différence est significative sans ou avec la correction de Holm.

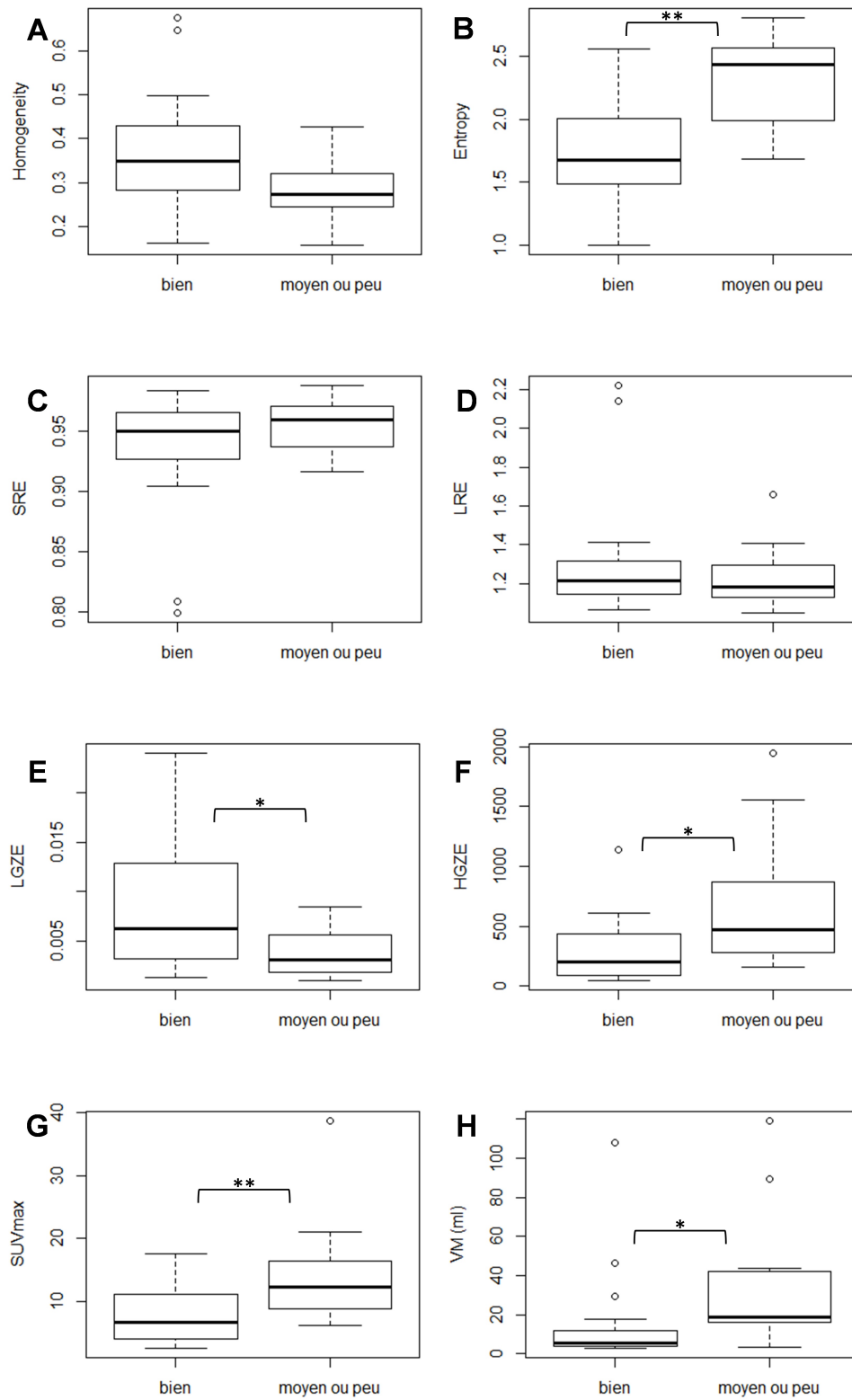


FIGURE 4.14 – Boîtes à moustache entre les lésions bien différenciées et les lésions moyennement ou peu différenciées (index de texture calculés entre 0 et 20 unités SUV). * et ** signifient respectivement que la différence est significative sans ou avec la correction de Holm.

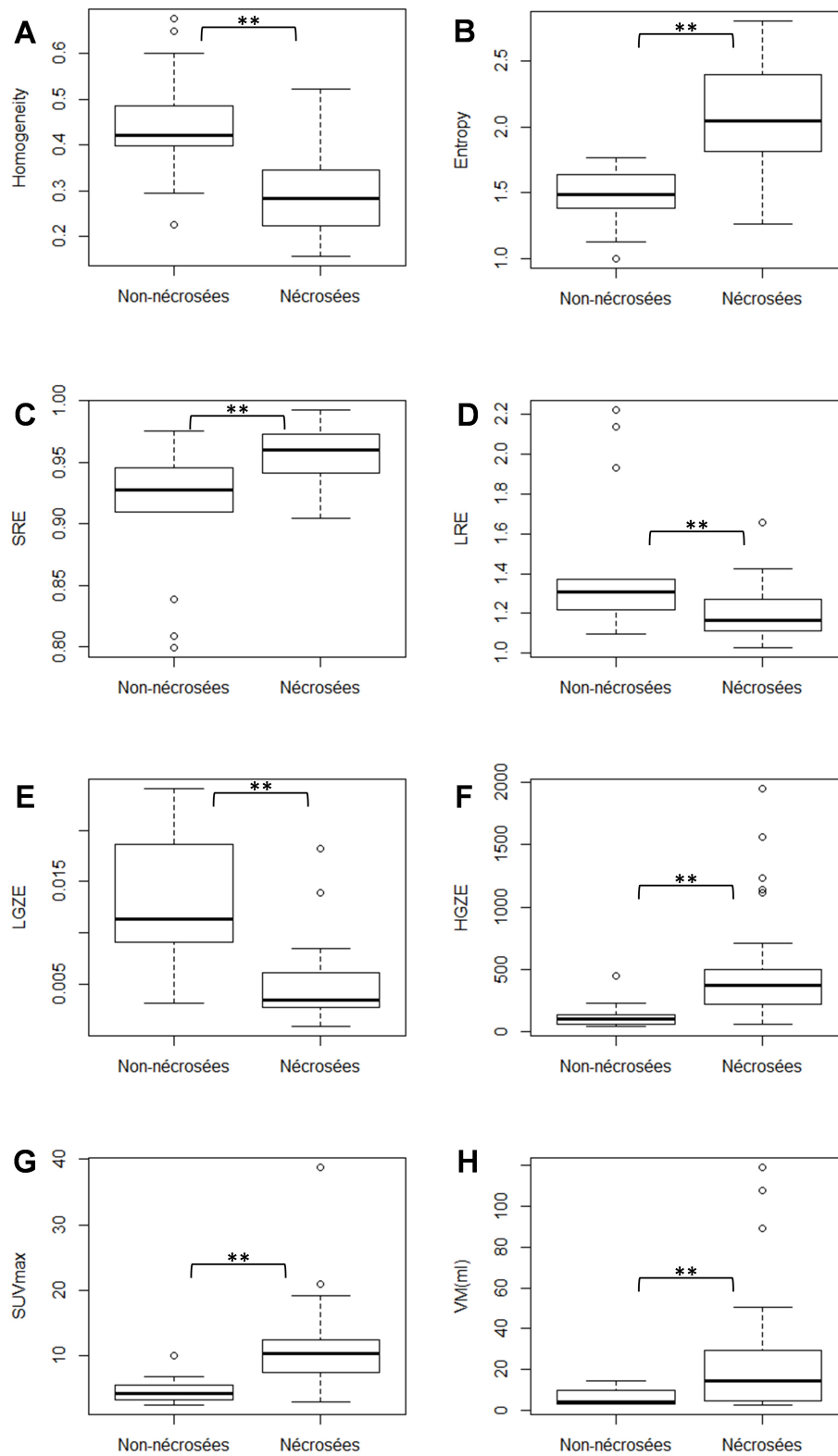


FIGURE 4.15 – Boîtes à moustache entre les lésions présentant ou non une nécrose histologique (index de texture calculés entre 0 et 20 unités SUV). * et ** signifient respectivement que la différence est significative sans ou avec la correction de Holm.

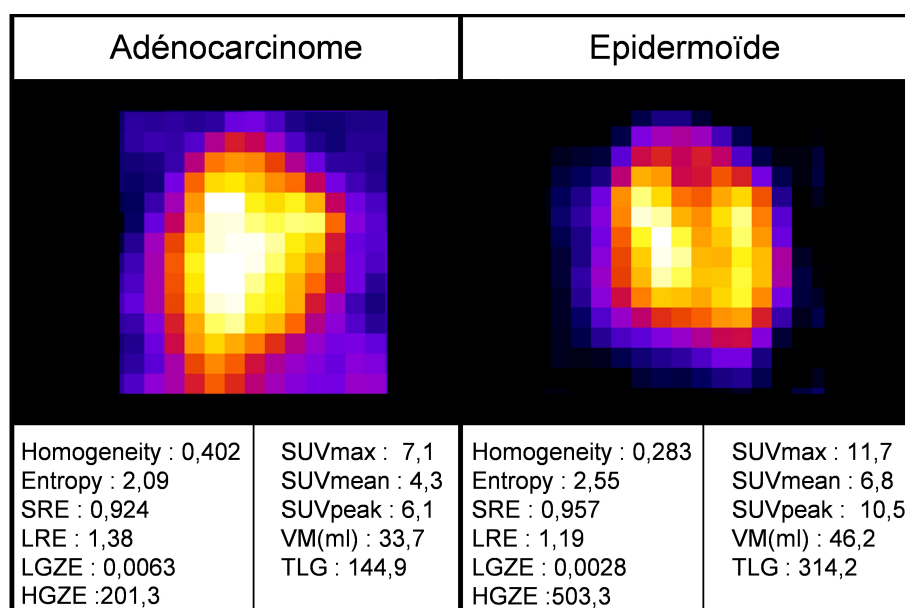


FIGURE 4.16 – Exemple d’une lésion adénocarcinome et d’une lésion épidermoïde avec les index de texture associés (calculés après DA entre 0 et 20 unités SUV).

4.3.2.3 Discussion

Comme pour les tumeurs mammaires, nos résultats suggèrent que les index de texture issus des images TEP sont sensibles à certaines caractéristiques des lésions pulmonaires.

Nous avons mis en évidence que VM et TLG étaient sensibles au stade tumoral et au stade ganglionnaire des lésions, mais pas Entropy lorsqu’on applique la correction de Holm. [Dong et al., 2013] ont montré une corrélation significative entre le stade tumoral ou ganglionnaire et Entropy chez 40 patients atteints d’un cancer de l’œsophage. Dans le cas de cancers pulmonaires non à petites cellules, [van Gómez López et al., 2014] ont rapporté une corrélation entre le stade tumoral et Entropy, VM et TLG sur 38 patients. Toutefois ces deux études n’ont ni rapporté le taux de discrétisation utilisé, ni le type de discrétisation employé et n’ont pas effectué de correction des comparaisons multiples.

Nos résultats montrent que les index de texture permettent de distinguer les deux sous-types de cancer avec des variations intuitives des paramètres. Par exemple, les adénocarcinomes sont plus homogènes que les épidermoïdes avec une plus faible valeur pour Entropy (à l’opposé de Homogeneity). Ces résultats sont cohérents avec ceux de [van Gómez López et al., 2014] qui montrent que les adénocarcinomes conduisent à un index Homogeneity plus élevé que les épidermoïdes. Dans leur étude, [Asamura et al., 2008] démontrent que les patients porteurs de cancers épidermoïdes ont un plus mauvais pronostic que les patients souffrant d’un adénocarcinome après chirurgie sur 11 939 patients. De plus, les épidermoïdes présentent un plus fort index de prolifération comparé aux adénocarcinomes [Mangili et al., 1998]. Ces résultats suggèrent que Homogeneity et Entropy pourraient permettre d’évaluer l’agressivité tumorale en combinant dans un même paramètre à la fois des informations relatives à l’hétérogénéité tumorale et à la fixation.

Les index de texture issus des images TEP sont également sensibles au niveau de différenciation des lésions tumorales et/ou à la présence ou non d’une nécrose histologique. Les variations des index

sont également cohérentes ici avec une texture plus hétérogène pour les lésions nécrosées. Or, la nécrose histologique a été montrée comme un facteur pronostique en oncologie [Richards et al., 2011], les résultats suggèrent donc que les paramètres de texture issus des images TEP pourraient fournir une information pronostique dans les cancers pulmonaires.

Parmi les index étudiés, Entropy conduit même à la plus faible p-value pour identifier le sous-type de cancer, le niveau de différenciation et la présence d'une nécrose histologique dans les lésions pulmonaires. Ce paramètre a déjà été rapporté comme un facteur pronostique dans les cancers pulmonaires non à petites cellules pour, d'une part, 45 patients [Pyka et al., 2015] et, d'autre part, 101 patients [Hatt et al., 2015]. Toutefois, ces deux publications utilisent une méthode de discrétisation relative qui augmente la corrélation entre les paramètres de texture et VM par un artefact de calcul (cf. paragraphe 3.4.3). Ces résultats nécessitent donc d'être confirmés avec une discrétisation absolue, une plus large cohorte et pour d'autres types de tumeurs.

4.3.3 Les gliomes

Dans cette troisième étude, nous avons cherché à déterminer si l'analyse de texture permettait de distinguer des tumeurs nouvellement diagnostiquées des tumeurs en récurrence, et d'identifier le grade tumoral dans les gliomes. Ce travail a été réalisé en collaboration avec Christophe Nioche (Hôpital d'instruction des armées du Val de Grâce).

4.3.3.1 Matériels & méthodes

Pour réaliser ce travail, les données de 81 patients porteurs de gliomes ont été étudiées. La constitution de ce groupe de patients à l'hôpital du Val de Grâce a été approuvée par le comité d'éthique local et un consentement écrit a été donné par les patients. Les caractéristiques de cette cohorte sont résumées dans le Tableau 4.11.

Caractéristiques	Lésions nouvellement diagnostiquées	Lésions en récurrence
Sexe		
Femme	17	12
Homme	35	17
Age (années)		
Moyenne	48±11	51±11
Intervalle	22-75	32-74
Histopathologie		
Bas Grade	23	8
<i>OMS grade I</i>	0	0
<i>OMS grade II</i>	23	8
Astrocytome gémostokystique	2	0
Oligoastrocytome	6	3
Oligodendrogliome	15	5
Haut Grade	29	21
<i>OMS grade III</i>	16	12
Astrocytome anaplasique	3	3
Oligoastrocytome anaplasique	5	4
Oligodendrogliome anaplasique	8	5
<i>OMS grade IV</i>	13	9
Glioblastome	13	9

Tableau 4.11 – Caractéristiques des patients porteurs de gliomes.

Pour chaque patient, l'acquisition dynamique des images TEP a été réalisée sur une machine Gemini temps de vol (Philips) durant 40 min, 3 min après l'injection de ^{18}F -LDOPA (2,2 MBq/kg), avec une position de lit centrée sur le cerveau (un champ de vue axiale de 25,6 cm). Une seule image TEP a été reconstruite correspondant à l'acquisition totale de 40 min. Les images ont été corrigées de l'atténuation à partir de l'acquisition TDM et aucun post-filtrage n'a été réalisé. Pour l'acquisition des images TDM, sans injection de produit de contraste, les caractéristiques utilisées étaient les suivantes : 120 kV, 224 mA et un pitch de 1,1. La taille des voxels dans les images reconstruites, à la fois en TEP et en TDM, était de 2 mm × 2 mm × 2 mm.

Les lésions ont été segmentées avec un seuillage adaptatif comme décrit précédemment (cf. section 3.1). Pour chaque lésion, cinq index conventionnels ont été calculés (SUVmax, SUVmean, SUVpeak, VM et TLG) ainsi que 6 paramètres de texture (Homogeneity, Entropy, SRE, LRE, HGZE, LGZE) après une étape de discrétisation absolue obtenue avec 64 niveaux de gris entre 0 et 20 unités SUV.

Pour les tumeurs de haut grade et de bas grade séparément, nous avons étudié la capacité des index à distinguer les lésions nouvellement diagnostiquées, des lésions en récurrence, à l'aide d'un test de Wilcoxon avec une correction de Holm pour les comparaisons multiples.

Nous avons ensuite testé si les index étaient capables de séparer les lésions de bas grade des lésions de haut grade parmi les lésions nouvellement diagnostiquées ($n=52$) et les lésions en récurrence ($n=29$). Pour ce faire, nous avons utilisé l'aire sous la courbe ROC (AUC) pour les index de texture seuls puis en combinaison avec SUVmax dans un modèle de régression binomial. Nous avons également calculé simultanément la p-value associée au test de Wilcoxon avec correction de Holm pour l'index de texture seul et en combinaison avec SUVmax.

4.3.3.2 Résultats

Différenciation entre les tumeurs nouvellement diagnostiquées et en récurrence

Parmi les 81 patients étudiés, 31 sont porteurs de tumeurs de bas grade, dont 23 sont des tumeurs nouvellement diagnostiquées et 8 sont des tumeurs en récurrence. Les cinquante autres patients sont porteurs de tumeurs de haut grade, dont 29 sont des patients nouvellement diagnostiqués (cf. Tableau 4.11).

L'analyse de texture des images TEP a montré que les index étaient capables de distinguer les tumeurs sans traitement des tumeurs en récurrence uniquement dans le groupe des lésions de haut grade (cf. Tableau 4.12). Les p-values des index de texture varient entre 0,0014 et 0,0110 et seuls Homogeneity et Entropy restent significatifs après la correction de Holm. La p-value de SUVmax est égale à 0,0023. À l'inverse, le VM n'est pas capable de faire la distinction entre les deux types de lésions. Comme le montre la Figure 4.18, les tumeurs en récurrence ont une texture en moyenne plus hétérogène que les tumeurs nouvellement diagnostiquées avec Homogeneity, LRE et LGZE plus faibles et Entropy, SRE, HGZE plus élevés.

Distinction du grade tumoral

Pour les tumeurs nouvellement diagnostiquées, les performances des index de texture pour séparer bas grade ($n=23$) et haut grade ($n=29$) sont résumées dans le Tableau 4.13. Entropy conduit à l'AUC la plus élevée ($AUC=0,70$) suivi des index Homogeneity, SRE et LRE. Pour la distinction entre les grades, les mesures SUV conduisent à des AUC entre 0,62 et 0,66 alors que le VM et le TLG conduisent à des AUC respectivement égales à 0,55 et 0,51. Quand on combine les index de texture avec SUVmax, le couple Entropy-SUVmax conduit à la plus forte AUC égale à 0,78 et permet donc d'améliorer la classification des lésions, même si l'AUC de Entropy-SUVmax n'est pas significativement supérieure à celle de SUVmax seul (d'après un test de Delong). Concernant les tests de Wilcoxon, aucune p-value n'apparaît significative après la correction de Holm, excepté la combinaison entre SUVmax et Entropy ou Homogeneity.

Index	Lésions nouvellement diagnostiquées vs. Lésions en récidive	
	Haut Grade	Bas Grade
Homogeneity	0,0045*	0,8814
Entropy	0,0014*	0,6234
SRE	0,0123	0,983
LRE	0,0138	0,9151
LGZE	0,0091	0,1354
HGZE	0,0110	0,2197
SUVmax	0,0023*	0,2927
SUVmean	0,0076	0,1354
SUVpeak	0,0023*	0,2927
VM	0,5327	0,5355
TLG	0,2568	0,2036

Tableau 4.12 – P-values du test de Wilcoxon entre les lésions nouvellement diagnostiquées et les lésions en récidive, pour les tumeurs de haut grade et de bas grade (index de texture calculés après DA entre 0 et 20 unités SUV). Les p-values en gras sont significatives au seuil de 5%. * signifie que la p-value reste significative après la correction de Holm.

Pour les lésions en récidive, les AUC des index de texture ($AUC=[0,85-0,90]$) sont plus faibles que celles des mesures SUV ($AUC=0,93$), et la combinaison des paramètres n'améliore pas la classification des lésions.

On observe toutefois que dans les deux groupes de patients, les variations des index de texture sont intuitives, avec une texture plus hétérogène des lésions de haut grade par rapport aux lésions de bas grade (cf. Figures 4.19 et 4.20). Par exemple sur les lésions présentées sur la Figure 4.17, pour la lésion de bas grade Homogeneity est égal à 0,854, Entropy à 0,63 contre 0,840 et 0,93 dans la lésion de haut grade.

4.3.3.3 Discussion

La capacité des index de texture à différencier des tumeurs nouvellement diagnostiquées des tumeurs en récidive suggère qu'il est possible de quantifier l'organisation tumorale à partir des images TEP et que ces paramètres peuvent être sensibles à des perturbations liées au traitement. Cette différenciation n'a été observée que pour les tumeurs de haut grade, avec une puissance statistique de l'ordre de 85%. Pour les tumeurs de bas grade, avec les effectifs actuels ($n=23$ lésions nouvellement diagnostiquées, $n=8$ lésions en récidive), la puissance statistique serait de 57% si l'on observait le même écart entre les deux groupes de lésions que dans les tumeurs de haut grade. Cette puissance est insuffisante pour mettre en évidence un effet. Il faudrait ajouter 10 lésions en récidive supplémentaires pour obtenir un rééquilibrage des groupes et une puissance statistique supérieure à 80%. Nous ne pouvons donc pas savoir si les images TEP ne reflètent par les différences de texture associées au caractère primaire ou en récidive des tumeurs de bas grade ou si nous n'avons pas pu mettre en évidence cette différenciation dû au trop faible effectif de nos échantillons.

Index	Lésions nouvellement diagnostiquées				Lésions en récidence			
	Index seul		SUVmax + Index		Index seul		SUVmax + Index	
	AUC [95% IC]	pvalue	AUC [95% IC]	pvalue	AUC [95% IC]	pvalue	AUC [95% IC]	pvalue
Homogeneity	0,69 [0,55-0,84]	0,0154	0,75 [0,61-0,89]	0,0016*	0,86 [0,68-1]	0,0019*	0,93 [0,84-1]	8,76e-05*
Entropy	0,70 [0,55-0,84]	0,0138	0,78 [0,65-0,91]	0,0004*	0,90 [0,77-2]	0,0004*	0,93 [0,82-1]	0,0001*
SRE	0,69 [0,54-0,83]	0,0188	0,70 [0,55-0,84]	0,0138	0,85 [0,66-1]	0,0028*	0,93 [0,82-1]	0,0001*
LRE	0,69 [0,55-0,84]	0,0154	0,70 [0,55-0,84]	0,0138	0,85 [0,66-1]	0,0034*	0,93 [0,84-1]	8,76e-05*
LGZE	0,56 [0,40-0,72]	0,4725	0,67 [0,52-0,82]	0,0319	0,86 [0,72-1]	0,0019*	0,93 [0,83-1]	0,0001*
HGZE	0,60 [0,45-0,76]	0,2032	0,66 [0,51-0,81]	0,0457	0,89 [0,76-1]	0,0008*	0,94 [0,84-1]	6,34e-05*
SUVmax	0,66 [0,51-0,81]	0,0457	-	-	0,93 [0,82-1]	0,0001*	-	-
SUVmean	0,62 [0,47-0,78]	0,1319	-	-	0,93 [0,82-1]	0,0001*	-	-
SUVpeak	0,66 [0,50-0,81]	0,0543	-	-	0,93 [0,84-1]	8,76e-05*	-	-
VM	0,55 [0,39-0,72]	0,5174	0,67 [0,52-0,82]	0,0366	0,54 [0,33-0,76]	0,7560	0,94 [0,86-1]	6,34e-05*
TLG	0,51 [0,35-0,67]	0,9083	0,66 [0,51-0,81]	0,0419	0,72 [0,54-0,90]	0,0745	0,95 [0,87-1]	4,47e-05*

Tableau 4.13 – Aire sous la courbe ROC [intervalle de confiance à 95%] et p-value du test de Wilcoxon pour les index seuls ou en combinaison avec SUVmax (régression logistique) pour la distinction entre les gliomes de bas grade et de haut grade pour les lésions nouvellement diagnostiquées ou en récidence. Les index de texture ont été calculés après une discrétisation absolue entre 0 et 20 unités SUV. Les p-values en gras sont significatives au seuil de 5%. * signifie que la p-value reste significative après la correction de Holm.

Les variations des index montrent une hétérogénéité et une fixation plus forte dans les lésions en récidence que dans les lésions nouvellement diagnostiquées. Un index de texture, Entropy, conduit à une p-value inférieure à celle de SUVmax, démontrant l'intérêt d'intégrer ces paramètres dans l'analyse des images et en particulier pour les tumeurs cérébrales qui n'ont fait l'objet, pour l'instant, d'aucune étude portant sur la texture à partir des images TEP.

Pour l'identification du grade tumoral, l'analyse des résultats montre que les paramètres de texture permettent d'améliorer la classification des tumeurs dans les lésions nouvellement diagnostiquées comparé à SUVmax seul. À l'inverse, cette combinaison n'apportent pas d'informations supplémentaires dans les lésions en récidence.

Les résultats obtenus pour les mesures SUV sont cohérents avec ceux présentés dans la littérature [Nioche et al., 2013]. En France, pour les patients atteints de gliomes, le traceur 18F-LDOPA est utilisé principalement pour la détection des récurrences. En effet, nous observons bien de meilleures

performances des mesures SUV pour les lésions en récidence que pour les lésions nouvellement diagnostiquées.

Nos observations suggèrent que les index de texture permettent d’améliorer la classification des gliomes nouvellement diagnostiqués en apportant aux mesures SUV une information relative à l’hétérogénéité tumorale. Une fois encore, c’est l’index Entropy qui semble le plus approprié pour compléter les informations issues des mesures SUV pour des lésions nouvellement diagnostiquées.

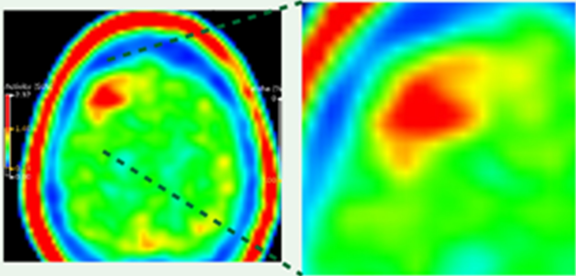
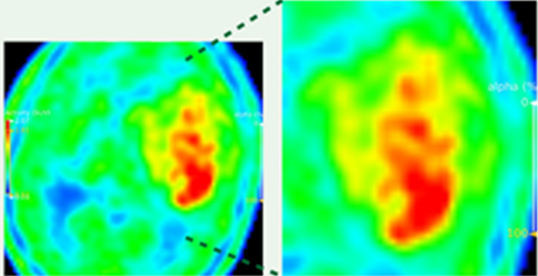
Lésion de bas grade		Lésion de haut grade	
			
Homogeneity : 0,854 Entropy : 0,63 SRE : 0,553 LRE : 8,28 LGZE : 0,0521 HGZE : 20,7	SUVmax : 1,6 SUVmean : 1,1 SUVpeak : 1,5 VM(ml) : 9,9 TLG : 10,9	Homogeneity : 0,840 Entropy : 0,93 SRE : 0,463 LRE : 12,3 LGZE : 0,0548 HGZE : 28,2	SUVmax : 2,1 SUVmean : 1,2 SUVpeak : 1,9 VM(ml) : 126,1 TLG : 151,3

FIGURE 4.17 – Exemple d’une lésion de bas grade et d’une lésion de haut grade, parmi les tumeurs nouvellement diagnostiquées, avec les valeurs des index de texture (calculés avec une discrétisation absolue entre 0 et 20 unités SUV).

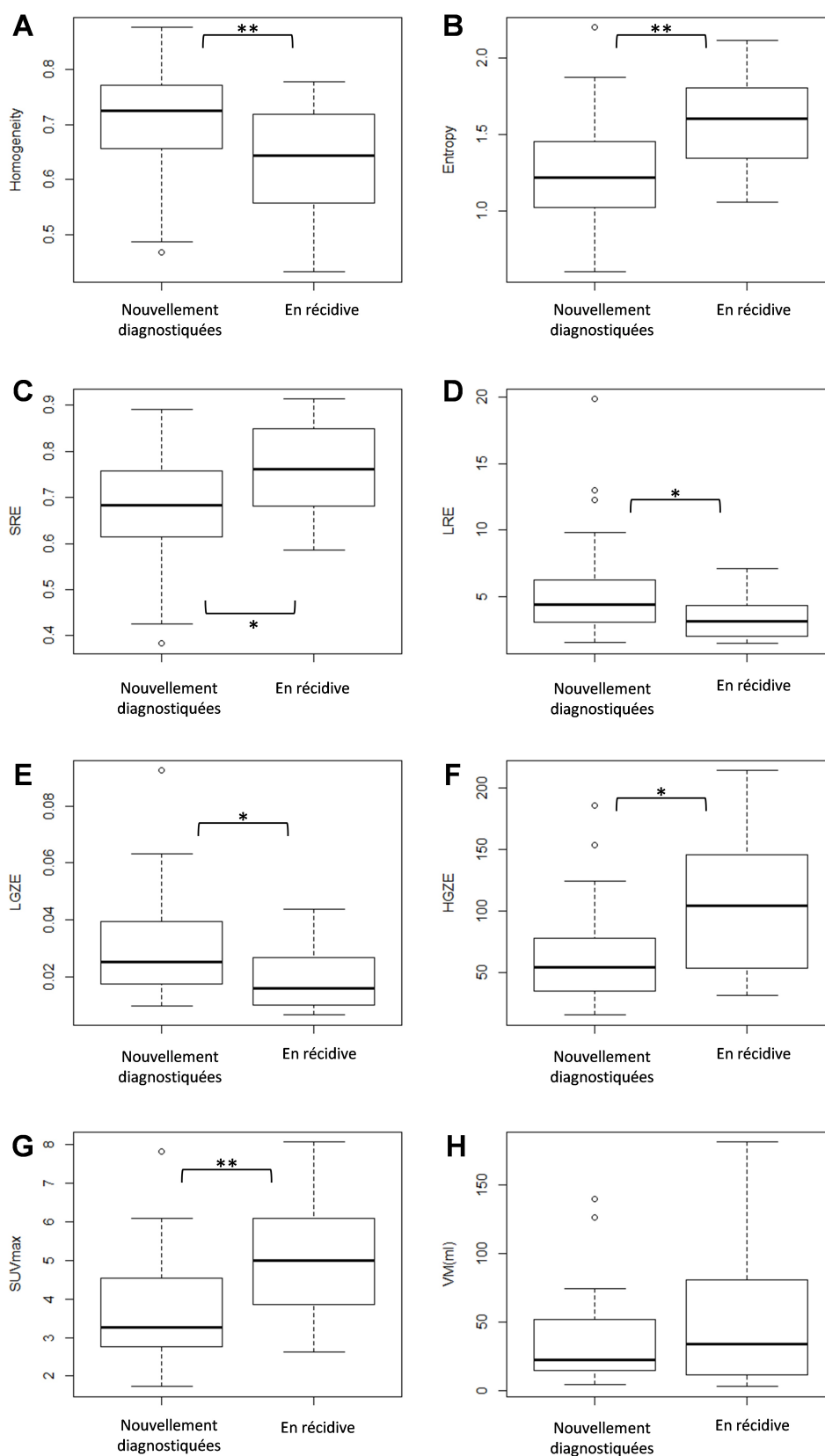


FIGURE 4.18 – Boîtes à moustache pour les index de texture, SUVmax et VM entre les lésions nouvellement diagnostiquées et les lésions en récive parmi les tumeurs de haut grade (index de texture calculés après DA entre 0 et 20 unités SUV). * et ** signifient respectivement que la différence est significative sans ou avec la correction de Holm.

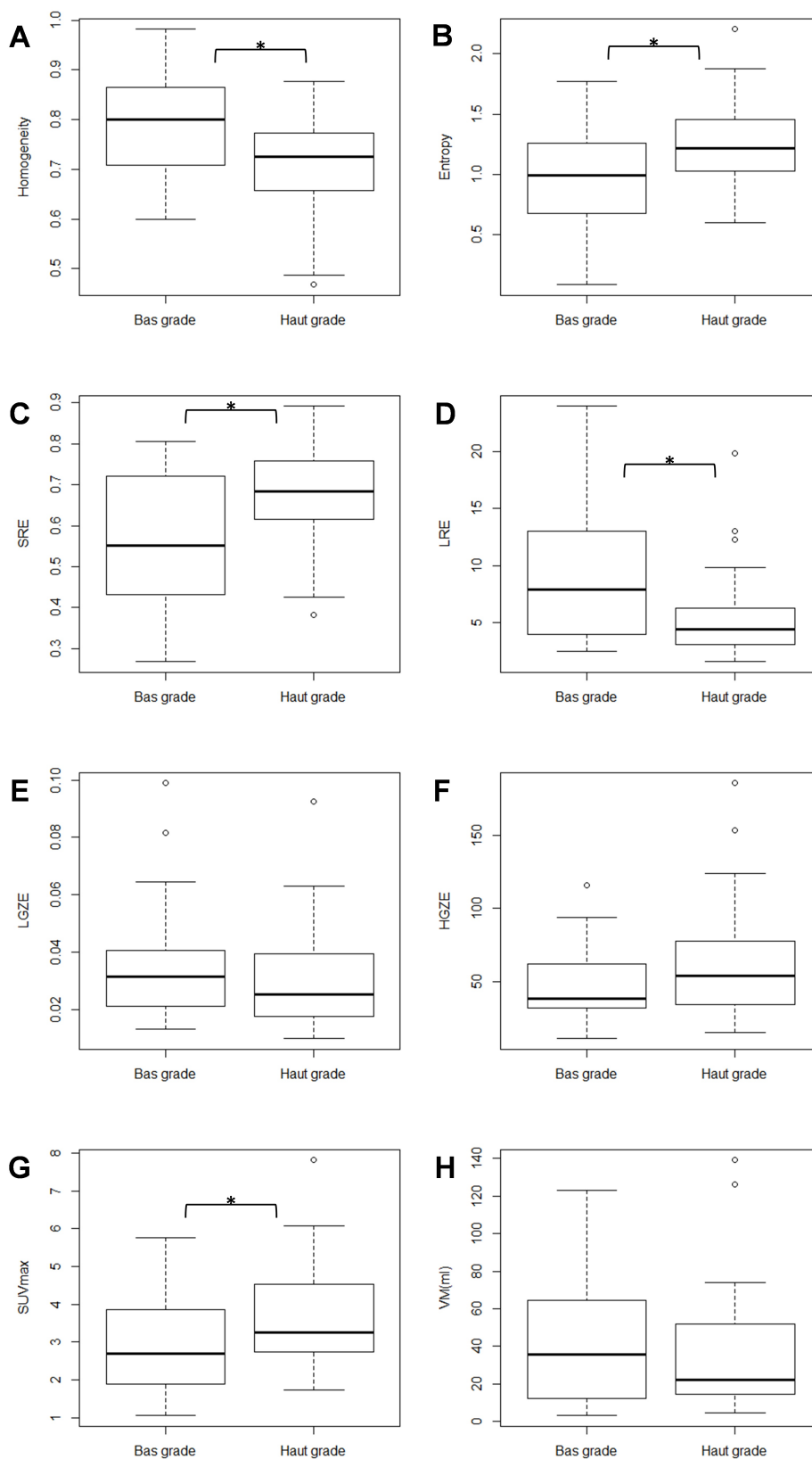


FIGURE 4.19 – Boîtes à moustache entre les lésions de bas grade et de haut grade parmi les lésions nouvellement diagnostiquées (index de texture calculés après DA entre 0 et 20 unités SUV). * et ** signifient respectivement que la différence est significative sans ou avec la correction de Holm.

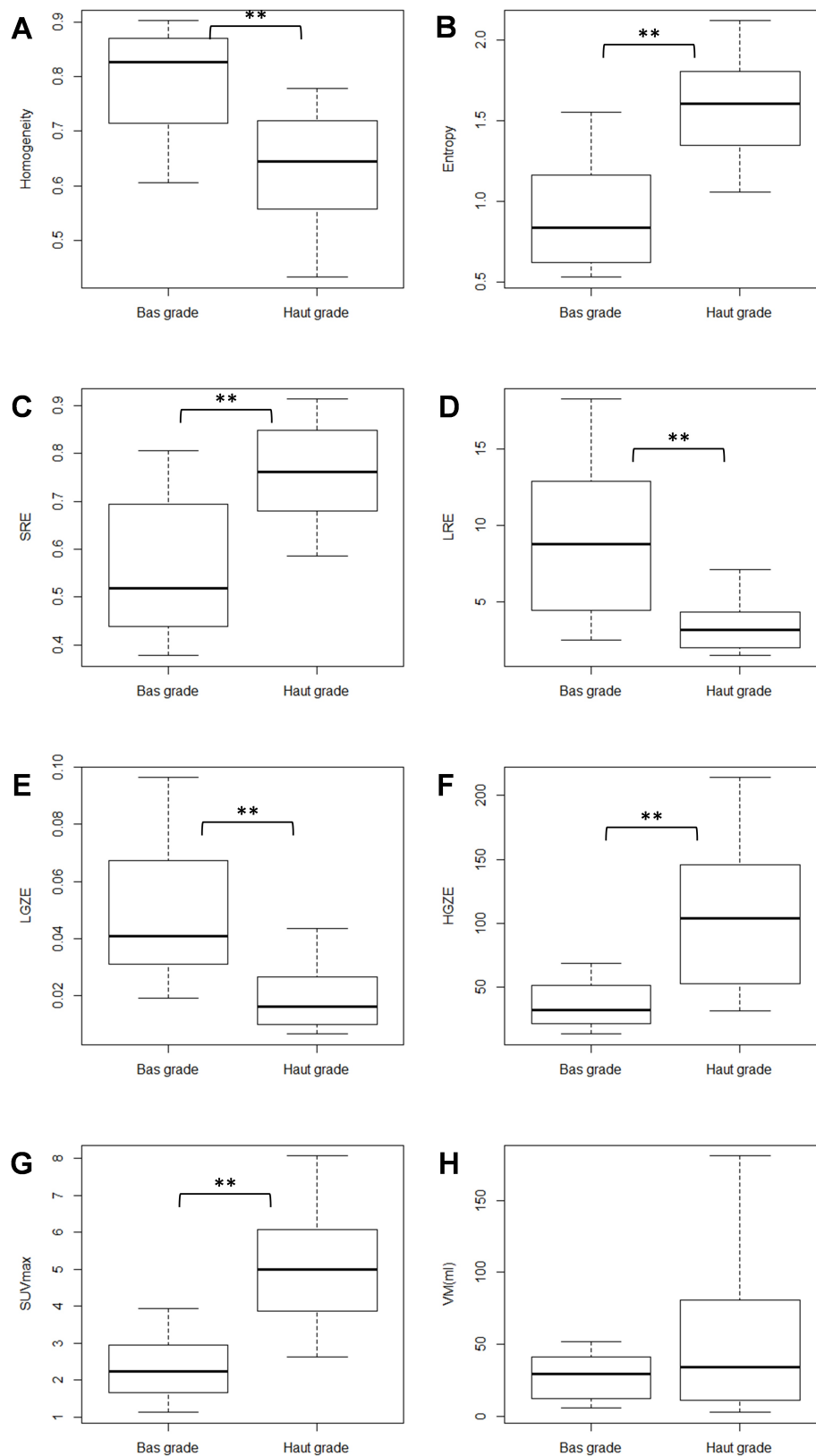


FIGURE 4.20 – Boîtes à moustache entre les lésions de bas grade et de haut grade parmi les lésions en récidive (index de texture calculés après DA entre 0 et 20 unités SUV). * et ** signifient respectivement que la différence est significative sans ou avec la correction de Holm.

Au cours de ce chapitre, nous avons montré que les paramètres de texture issus des images TEP étaient sensibles à la différence de texture entre les tissus sains et les tissus tumoraux mais également aux caractéristiques tumorales, grâce à un calcul approprié des index. Contrairement à ce qui avait été rapporté par [Brooks and Grigsby, 2015] avec l'utilisation de la discrétisation relative, les index permettraient de capturer une information pertinente pour l'évaluation tumorale. Dans nos études, parmi les six index de texture calculés, Entropy conduit le plus souvent aux meilleures performances même si ces conclusions nécessitent d'être validées sur de plus grandes cohortes, pour d'autres types de cancers et d'autres caractéristiques tumorales.

Dans le chapitre suivant, nous chercherons à comprendre davantage la signification biologique de ces paramètres en comparant l'hétérogénéité mesurée à différentes échelles.

Chapitre 5

Étude multi-échelle de l'hétérogénéité tumorale

Après avoir effectué une analyse méthodologique et clinique des paramètres de texture, nous avons pu établir un protocole de calcul strict et démontrer l'intérêt de ces index pour caractériser les lésions tumorales à partir des images TEP. Dans la littérature, certains auteurs ont montré que l'analyse de texture permettait de prédire la réponse aux traitements ou la survie du patient dans plusieurs types de cancers (cf. Chapitre 2) mais sans étudier la relation entre les index de texture issus des images TEP et l'organisation cellulaire au niveau des coupes histologiques. C'est l'objectif de ce dernier chapitre : comprendre quel est le lien entre l'hétérogénéité tumorale mesurée à différentes échelles et quelles informations peuvent être reflétées par une analyse de texture des images TEP.

5.1 Matériels & méthodes

Les modèles animaux utilisés ainsi que les protocoles d'acquisition et de traitement des images sont détaillés dans cette partie.

5.1.1 Modèles animaux

Cette étude a été approuvée par le Comité d'ETHique en Expérimentation Animale (CETEA DSV n°44) sous la référence 12-036 et réalisée conformément à la réglementation européenne portant sur les animaux de laboratoire. Trois souris FVB, greffées orthotopiquement à l'aide de cellules tumorales mammaires dérivées de souris transgéniques MMTV-PyMT, ont été utilisées pour cette étude.

5.1.2 Protocole de l'imagerie TEP/TDM

Trente à trente-six jours après l'implantation des tumeurs, les animaux ont été mis à jeun 12 heures avant la réalisation de l'examen TEP et ont été anesthésiés en utilisant 2% d'isoflurane gazeux. L'examen TEP/TDM a été réalisé sur une machine Inveon TEP/TDM (Siemens) dédiée à l'imagerie du petit animal. Chaque souris a été injectée avec $7,9 \pm 0,3$ MBq de ^{18}F -FDG, 65 \pm 5 min avant le début de l'acquisition qui a duré 30 min. La fenêtre en énergie a été fixée entre 350 et 650 keV.

L'image TEP a été reconstruite en utilisant un algorithme OSEM en 2D avec 4 itérations et 16 sous-ensembles. La correction de la diffusion a été réalisée directement à partir du calcul des formules analytiques [Watson, 2000] et la correction de l'atténuation à partir des images TDM. Les images n'ont pas été post-filtrées. La taille de la matrice est de $256 \times 256 \times 159$ voxels et la taille des voxels sur les images TEP reconstruites est de $388 \times 388 \times 796 \mu\text{m}^3$.

5.1.3 Autoradiographie

Après l'acquisition TEP/TDM, les souris ont été sacrifiées et les tumeurs ont été prélevées. Chaque tumeur a ensuite été détaillée à l'aide d'un cryostat (Leica) en coupe de $20\ \mu\text{m}$ d'épaisseur tous les $100\ \mu\text{m}$ et déposée sur une lame de verre. Pour la première tumeur, nous avons obtenu 116 coupes et pour les deux autres 102 coupes.

Les coupes sont alors placées sur des plaques radiosensibles pendant environ 15 heures avant d'être numérisées en utilisant un scanner de type Storm (GE Healthcare). La taille des voxels sur l'image numérisée est de $50 \times 50 \times 20\ \mu\text{m}^3$.

5.1.4 Histochemie

Pour réaliser une analyse histochemie, les coupes ont été colorées à l'hématoxyline et à l'éosine (H&E). Ce marquage permet la coloration en bleu-violet du noyau des cellules et en rose du cytoplasme et de la matrice extracellulaire.

Pour séparer la composante hématoxyline et obtenir une distribution des cellules, nous avons utilisé le module *Color Deconvolution* [Ruifrok and Johnston, 2001] du logiciel ImageJ (cf. Figure 5.1). Sur la composante « violette », nous avons ensuite lissé les images avec un filtre gaussien (écart-type = $0,48\ \mu\text{m}$) et ré-échantillonné pour obtenir la même taille de voxels que sur les images autoradiographiques ($50 \times 50 \times 20\ \mu\text{m}^3$) à l'aide du module *TransformJSscale* [Meijering et al., 2001] de ImageJ. Ces images seront nommées H dans la suite. Un second lissage a ensuite été réalisé sur les images H avec un filtre gaussien (écart-type = $300\ \mu\text{m}$) permettant d'obtenir des images appelées HL, afin de se rapprocher de la résolution spatiale des images autoradiographiques.

5.1.5 Préparation des volumes 3D

5.1.5.1 Construction, recalage, segmentation

Dans un premier temps, une fois les images autoradiographiques numérisées, un volume 3D a été reconstruit avec le module *StrackReg* [Thevenaz et al., 1998] (transformation rigide) de ImageJ. L'intervalle manquant de $80\ \mu\text{m}$ entre deux coupes a été reconstitué par une transformation de type cubique β -spline (*TransformJSscale*) en faisant l'hypothèse que les changements morphologiques sont négligeables sur une si petite distance. La taille des voxels dans le volume 3D reconstruit est de $50 \times 50 \times 20\ \mu\text{m}^3$. Nous avons alors reconstitué trois volumes 3D (cf. Figure 5.1) :

- un volume à partir des coupes autoradiographiques d'origine (AR, voxel= $50 \times 50 \times 20\ \mu\text{m}^3$);
- un volume à partir des coupes autoradiographiques ré-échantillonnées pour avoir la même taille de voxels que sur les images TEP (ARR, $388 \times 388 \times 796\ \mu\text{m}^3$) en utilisant le module *TransformJSscale*;
- un volume à partir des coupes autoradiographiques ré-échantillonnées et lissées pour avoir la même taille de voxels et la même résolution spatiale que sur les images TEP (ARRL, $388 \times 388 \times 796\ \mu\text{m}^3$) en utilisant le module *TransformJSscale* et un filtrage gaussien (écart-type = $677\ \mu\text{m}$).

Dans un second temps, le volume TEP est recalé sur le volume ARRL (cf. Figure 5.1) en utilisant le module *Optimized Automatic Image Registration* du logiciel MIPAV (*Medical Image Processing Analysis and Visualization*) [McAuliffe et al., 2001], grâce à une interpolation tri-linéaire avec 12 degrés de liberté et minimisant une fonction de rapport de corrélation.

Après la reconstruction des volumes TEP, AR, ARR et ARRL, une segmentation de la tumeur a été réalisée, indépendamment sur chaque volume, en appliquant un seuillage fixe égal à 40% de la fixation maximale dans le volume.

Pour les images issues de l'analyse histochimique, H et HL, nous avons reconstruit les volumes 3D avec le module *StrackReg* (transformation rigide) et segmenté manuellement la tumeur du fond.

5.1.5.2 Normalisation

Les volumes TEP sont convertis en unités SUV en normalisant par la masse de l'animal comme présenté au paragraphe 1.2.1.2.1.

Pour les volumes AR, ARR et ARRL, une normalisation d'après les formules suivantes a été effectuée :

$$I(SUV) = \frac{\text{masse}(g) \times K^*(kBq) \times I(ua)}{Dose(t = t_{auto})} \quad (5.1)$$

$$K^*(kBq) = \frac{C_{TEP}(\frac{kBq}{mL}) \times V_{TEP}(mL)}{C_{auto}(\frac{ua}{mL}) \times V_{auto}(mL)} \quad (5.2)$$

où $I(ua)$ correspond à la valeur du voxel avant normalisation et $I(SUV)$ après normalisation. $Dose$ représente la dose injectée corrigée de la décroissance radioactive au moment de l'autoradiographie. K^* est égal au rapport entre la fixation moyenne dans le volume TEP après segmentation (C_{TEP}) multipliée par le volume segmenté en TEP (V_{TEP}) et la fixation moyenne dans le volume segmenté en autoradiographie (C_{auto}) multipliée par le volume segmenté en autoradiographie (V_{auto}). V_{TEP} et V_{auto} ont été obtenus suivant les procédures de segmentation décrites au paragraphe 5.1.5.1.

Cette étape de normalisation permet d'exprimer les images autoradiographiques en unités SUV afin de pouvoir les comparer directement aux images TEP. L'activité totale dans les images autoradiographiques normalisées par K^* est identique à celle dans les images TEP (équation 5.2). Elle est ensuite corrigée des doses injectées, des temps d'acquisition post-injection et des masses des souris qui sont différents entre les trois animaux (équation 5.1).

Pour les volumes H et HL, les images ont été normalisées pour chaque souris en utilisant un facteur de calibration égal à :

$$I(SUV) = K^{**}(SUV) \times I(ua) \quad (5.3)$$

$$K^{**}(SUV) = \frac{C_{auto}(\frac{SUV}{mL}) \times V_{auto}(mL)}{C_H(\frac{ua}{mL}) \times V_H(mL)} \quad (5.4)$$

où $I(ua)$ correspond à la valeur du voxel avant normalisation et $I(SUV)$ après normalisation. C_H correspond à la fixation moyenne dans le volume segmenté sur l'ensemble des coupes marquées à l'hématoxyline (H) multipliée par le volume segmenté sur l'ensemble des images H (V_H), obtenu suivant la procédure décrite au paragraphe 5.1.5.1.

Cette étape de calibration permet de normaliser en unités SUV les intensités mesurées sur les coupes marquées par rapport à la fixation mesurée en autoradiographie indépendamment sur les trois souris.

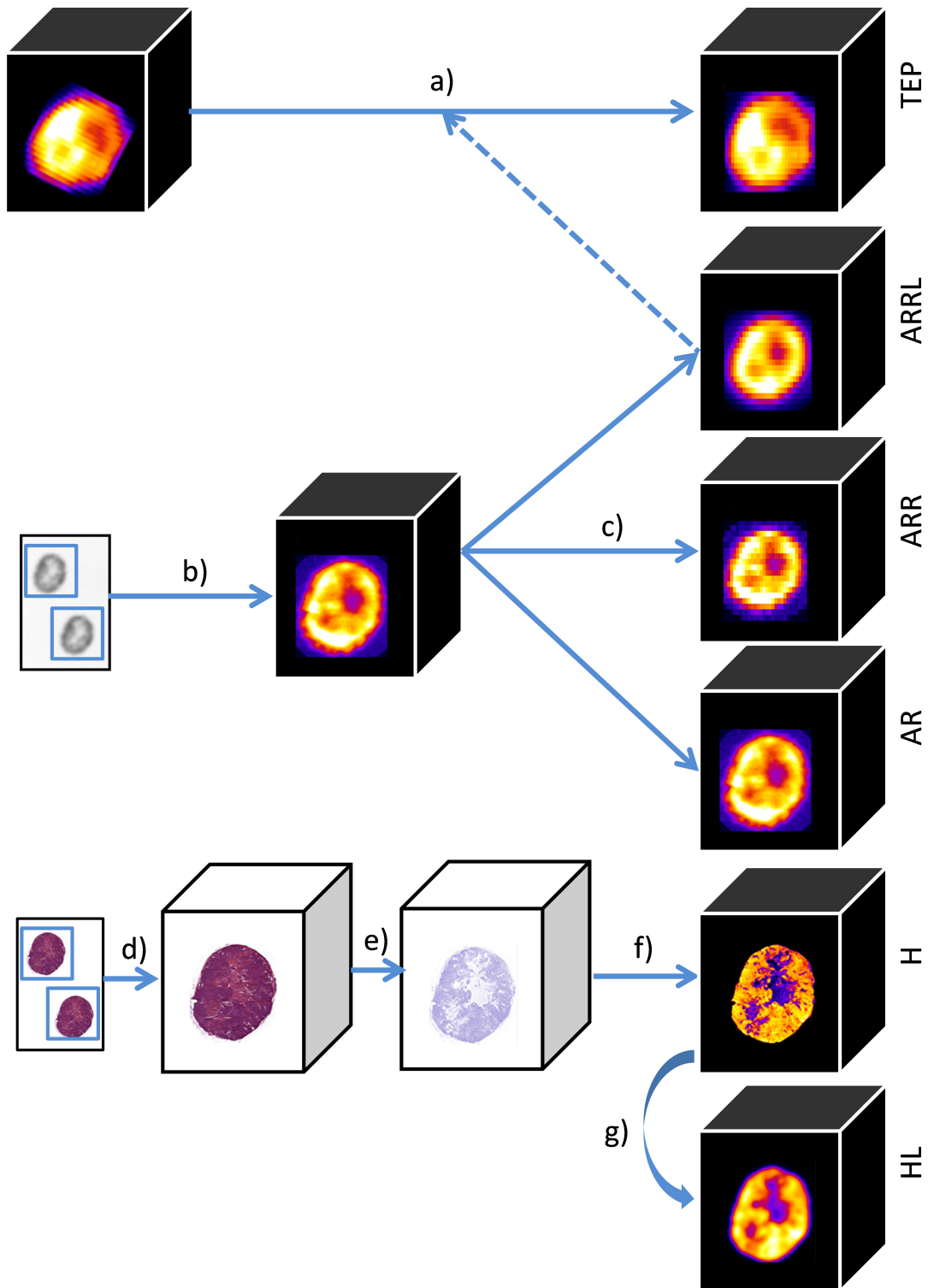


FIGURE 5.1 – Schéma de construction des volumes analysés. a) Recalage du volume TEP initial sur le volume ARRL. b) Après numérisation de la plaque autoradiographique, découpage des coupes puis recalage pour former un volume 3D. c) À partir de ce volume 3D nommé AR, création de deux autres volumes : ARR et ARRL. d) Découpage des coupes après marquage H&E puis recalage pour former un volume 3D. e) Déconvolution pour extraire la composante hématoxyline. f) Ré-échantillonnage et lissage pour obtenir le volume H. g) Lissage de H pour obtenir le volume HL.

5.1.6 Analyse des images

Des coupes ont été sélectionnées conjointement sur tous les volumes. Pour chacune de ces images en 2 dimensions et pour chaque modalité (TEP, AR, ARR, ARRL, H et HL), l'intensité maximale a été calculée (Max). Avant de construire les matrices de texture, nous avons discrétisé les intensités des voxels en utilisant une méthode de discrétisation absolue avec 64 niveaux de gris. La borne minimale pour cette discrétisation a été fixée à 0. La borne maximale correspond au maximum des Max sur l'ensemble des coupes de chaque modalité soit 6 pour les images TEP, 0,1 pour AR, 12 pour ARR, 9 pour ARRL, et 0,1 pour H et HL.

Sur chaque coupe et pour chaque modalité, nous avons calculé les six index de texture précédemment présentés : Homogeneity, Entropy, SRE, LRE, LGZE et HGZE.

5.2 Impact de la taille des voxels et de la résolution spatiale

Dans cette première étude, à partir de coupes acquises par autoradiographie, le but a été de comparer la texture sur des images provenant du même signal mais avec des tailles de voxels et des résolutions spatiales différentes.

5.2.1 Matériels & méthodes

Nous avons comparé les index de texture extraits des trois types d'autoradiographie décrits précédemment. Sur les coupes sélectionnées, nous avons calculé le coefficient de corrélation de Spearman (r_s) entre les index extraits de AR et de ARR et entre ceux extraits de ARR et ARRL afin d'étudier le rôle respectif de la taille des voxels et de la résolution spatiale. Conjointement, pour les mêmes paires d'index, nous avons effectué un test de Wilcoxon avec une correction de Holm pour vérifier si les valeurs absolues des paramètres de texture étaient influencées par la taille des voxels et la résolution spatiale.

5.2.2 Résultats

Pour la première souris, le volume tumoral était de $1,13 \text{ cm}^3$ ($2,26 \times 10^7$ voxels sur les images AR ou 9430 voxels sur les images ARR et ARRL). Pour la seconde souris, le volume tumoral correspondait à $1,39 \text{ cm}^3$ ($2,78 \times 10^7$ voxels sur les images AR ou 11599 voxels sur les images ARR et ARRL). Enfin, pour la troisième souris, le volume tumoral était égal à $1,20 \text{ cm}^3$ ($2,40 \times 10^7$ voxels sur les images AR ou 10014 voxels sur les images ARR et ARRL).

Sur les trois tumeurs utilisées pour cette étude, 28 coupes ont été sélectionnées avec un écart minimal entre deux coupes venant du même animal égal à $388 \mu\text{m}$.

Les résultats des coefficients de Spearman et des tests de Wilcoxon sont résumés dans le Tableau 5.1. Pour la comparaison AR vs. ARR, tous les index de texture ainsi que Max sont modérément ou fortement corrélés avec r_s supérieur ou égal à 0,62 (p-values de $r_s < 5\%$). Néanmoins, toutes les p-values du test de Wilcoxon sont largement inférieures à 5%, comme le confirme la Figure 5.2. On remarque que les images AR apparaissent plus homogènes que les images ARR avec des valeurs pour Homogeneity et Entropy, respectivement, plus élevées et plus faibles. HGZE et Max sont plus élevés sur les images ARR que sur AR. L'usage de voxels plus gros a donc introduit de l'hétérogénéité dans les images.

Concernant l'influence de la résolution spatiale, les index de texture sont également modérément ou fortement corrélés avec r_s compris entre 0,60 et 0,96. Homogeneity, Entropy, SRE et LRE ne sont pas significativement différents entre les images ARR et ARRL (test de Wilcoxon) à la différence de LGZE, HGZE et Max. La Figure 5.2 illustre ce constat avec des valeurs pour Homogeneity et Entropy proches de la première bissectrice entre la texture mesurée sur ARR et sur ARRL. À l'inverse, HGZE est plus élevé pour les images ARRL et Max plus élevé pour les images ARR.

Index	AR vs. ARR		ARR vs. ARRL	
	r_s	pvalue	r_s	pvalue
Homogeneity	0,71†	2,62e-16*	0,69†	0,1346
Entropy	0,77†	1,98e-10*	0,96†	0,9676
SRE	0,65†	2,62e-16*	0,60†	0,5310
LRE	0,62†	2,62e-16*	0,62†	0,0781
LGZE	0,92†	0,0012*	0,66†	0,0001*
HGZE	0,78†	9,13e-05*	0,59†	0,0002*
Max	0,89†	2,62e-16*	0,68†	0,0005*

Tableau 5.1 – Coefficients de corrélation de Spearman (r_s) et p-values du test de Wilcoxon pour l'étude de l'impact de la taille des voxels et de la résolution spatiale sur six index de texture et Max. † signifie que la corrélation est significative au seuil de 5%. Les p-values en gras sont significatives au seuil de 5%. * signifie que la p-value reste significative après correction de Holm.

5.2.3 Discussion

Nous avons comparé les valeurs des index de texture mesurés *ex vivo* à partir de coupes autoradiographiques afin d'étudier séparément l'effet de la taille des voxels et de la résolution spatiale sur les paramètres de texture. Nos résultats montrent que tous les index sont significativement corrélés entre AR et ARR et entre ARR et ARRL ($r_s \geq 0,60$). Néanmoins, les valeurs absolues de tous les index sont influencées par la différence de taille de voxels. Toutefois, cet effet est majoré ici (voxel = $50 \times 50 \times 20 \mu m^3$ contre $388 \times 388 \times 796 \mu m^3$) et ne correspond pas à ce qui serait observé en clinique entre différents imageurs TEP pour lesquels les différences de taille de voxels sont moins importantes.

L'index Entropy apparaît être le plus robuste vis-à-vis des différences en termes de résolution spatiale avec un coefficient de Spearman égal à 0,96 et une stabilité de l'index (test de Wilcoxon non-significatif) entre les images ARR et ARRL. LGZE, HGZE et Max semblent assez sensibles à la résolution spatiale avec une corrélation modérée entre la texture des deux images ($r_s = [0,59 - 0,68]$) et un écart sur les valeurs absolues (p-value du test de Wilcoxon < 5%).

L'étude de la corrélation entre AR vs. ARR et ARR vs. ARRL a également été réalisée sur l'ensemble des index calculés au Chapitre 3 (cf. Tableau 5.2). Les résultats obtenus pour un index par groupe de corrélation sont cohérents avec ceux observés pour les autres index du groupe. Par exemple, Dissimilarity qui appartient au groupe n°1, conduit à un r_s égal à 0,72 entre AR et ARR et 0,65 entre ARR et ARRL contre 0,71 et 0,69 respectivement pour Homogeneity (l'index que nous avons sélectionné pour représenter le groupe n°1).

La texture mesurée sur les images AR est plus homogène que celle mesurée sur les images ARR. Ce résultat correspond à l'appréciation visuelle (cf. Figure 5.1). Ce phénomène peut s'expliquer, dans

notre cas, par la présence de grandes plages homogènes, mais qui, étant différentes entre elles, rendent la tumeur hétérogène. L'utilisation de petits voxels fait alors apparaître de nombreuses régions homogènes à l'intérieur du volume tumoral. A l'inverse, l'utilisation de plus gros voxels entraîne une pixellisation et accentue les variations locales, sachant qu'avec de gros voxels, il y a moins de plages homogènes.

Dans la comparaison entre les images ARR et ARRL, le lissage modifie assez peu le comportement des index de texture. En pratique, le lissage des images permet de mieux visualiser l'hétérogénéité « vraie » en réduisant l'effet confondant du bruit. Or, lors du calcul des index de texture, l'étape de discrétisation permet déjà de réduire l'impact du bruit en regroupant des voxels d'intensités proches, la discrétisation agit donc elle-même un peu comme un lissage. Ce mécanisme peut expliquer la faible sensibilité des index de texture aux variations de résolution spatiale, quand ces variations restent restant modestes, et que le lissage ne gomme pas les hétérogénéités « macroscopiques » présentes dans la tumeur.

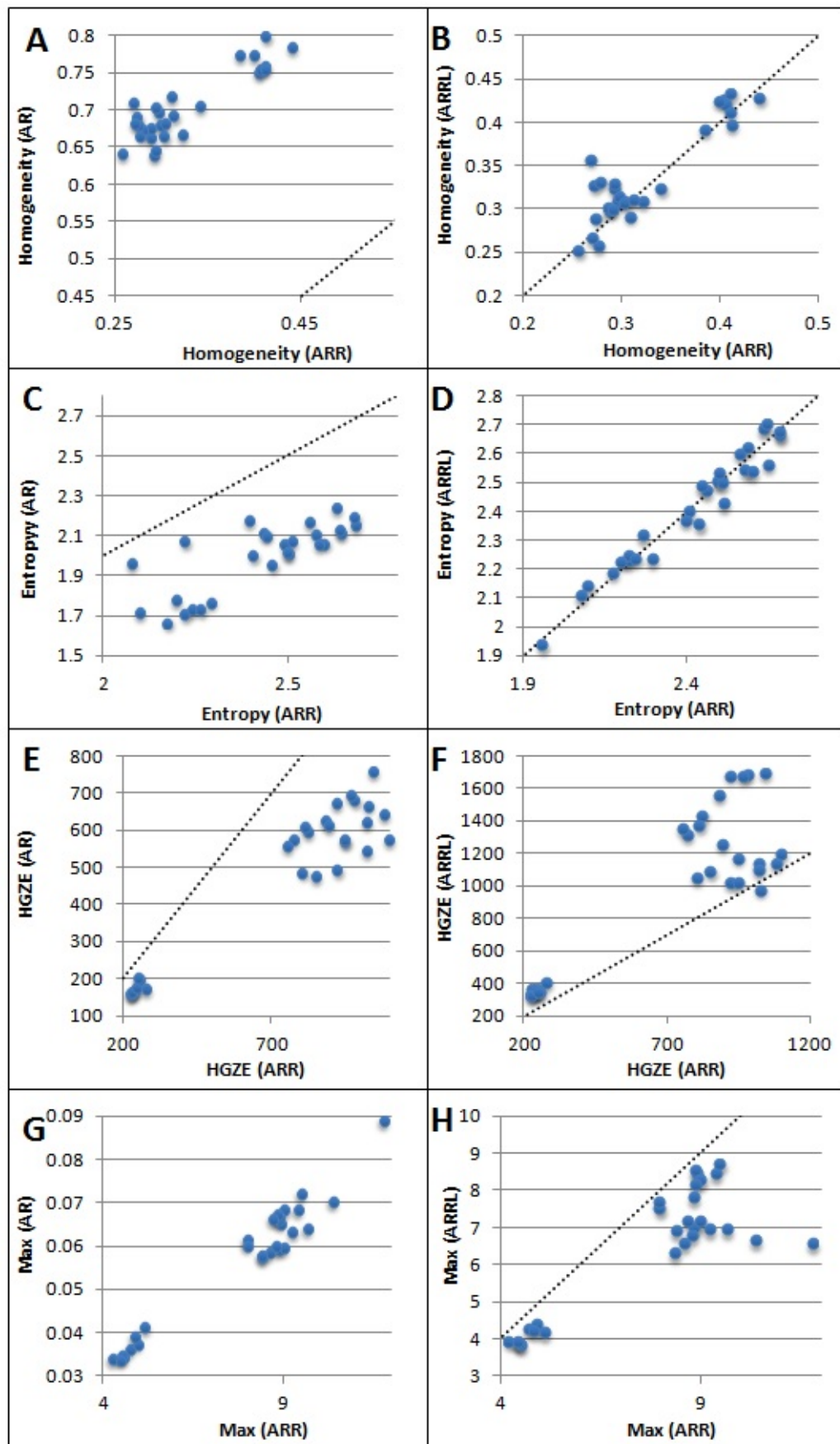


FIGURE 5.2 – Graphiques des index entre AR vs. ARR et ARR vs. ARRL. Les lignes en pointillés correspondent à la première bissectrice.

Index	AR vs. ARR		ARR vs. ARRL	
	r_s	pvalue	R	pvalue
Homogeneity	0,71†	2,62e-16*	0,69†	0,1346
Correlation	0,22	2,62e-16*	0,72†	8,66e-06*
Contrast (CM)	0,76†	2,62e-16*	0,62†	0,0214
Dissimilarity	0,72†	2,62e-16*	0,65†	0,0317
Energy	0,85†	4,63e-11*	0,95†	0,9535
Entropy	0,77†	1,98e-10*	0,96†	0,9676
SRE	0,65†	2,62e-16*	0,60†	0,5310
RP	0,64†	2,62e-16*	0,61†	0,2369
SZE	-0,61†	2,62e-16*	0,67†	0,3659
ZP	0,65†	2,62e-16*	0,62†	0,1143
LRE	0,62†	2,62e-16*	0,62†	0,0781
LZE	0,65†	2,62e-16*	0,62†	0,0753
LGRE	0,95†	0,0037*	0,70†	0,0002*
SRLGE	0,95†	0,9417	0,70†	0,0002*
LRLGE	0,83†	7,34e-07*	0,74†	0,0005*
LGZE	0,92†	0,0012*	0,66†	0,0001*
SZLGE	0,91†	0,0700	0,60†	0,0001*
HGRE	0,82†	9,86e-05*	0,63†	0,0003*
SRHGE	0,79†	8,45e-05*	0,62†	0,0004*
LRHGE	0,81†	7,10e-13*	0,68†	0,0001*
HGZE	0,78†	9,13e-05*	0,59†	0,0002*
SZHGE	0,69†	4,84e-05*	0,54†	0,0008*
GLNUr	0,90†	2,62e-16*	0,93†	0,9676
RLNU	0,86†	2,62e-16*	0,98†	0,6084
GLNUz	0,30	2,62e-16*	0,92†	0,6784
ZLNU	0,90†	1,79e-13*	0,97†	0,9935
Contrast (NGLDM)	0,93†	0,8645	0,97†	0,7636
Coarseness	0,91†	0,0517	0,96†	0,2241
Busyness	0,43†	0,3162	-0,26	0,0071
LZLGE	0,69†	2,62e-16*	0,74†	0,0060
LZHGE	-0,49†	2,62e-16*	0,72†	1,24e-06*

Tableau 5.2 – Coefficients de corrélation de Spearman (r_s) et p-values du test de Wilcoxon pour l'étude de l'impact de la taille des voxels et de la résolution spatiale sur tous les index de texture. † signifie que la corrélation est significative au seuil de 5%. Les p-values en gras sont significatives au seuil de 5%. * signifie que la p-value reste significative après correction de Holm.

5.3 Comparaison de l'analyse de texture à différentes échelles

Après avoir étudié l'impact de la résolution spatiale et de la taille des voxels pour des images issues d'une même acquisition, nous avons souhaité comparer la texture mesurée à différentes échelles : *in vivo* sur les images TEP, *ex vivo* sur les trois types d'autoradiographies et sur les coupes histologiques après marquage à l'hémaroxyline.

5.3.1 Matériels & méthodes

Pour réaliser cette analyse, nous avons calculé le coefficient de corrélation de Spearman entre la texture mesurée sur les images TEP et celle mesurée sur les images autoradiographiques (AR ou ARR ou ARRL) sur les 28 coupes sélectionnées. Nous avons également calculé r_s entre la texture issue des coupes après marquage à l'hémaroxyline (H et HL) et celle issue des images TEP ou AR. Parallèlement, un test de Wilcoxon avec une correction de Holm a été réalisé pour les mêmes comparaisons afin de vérifier la stabilité des valeurs absolues des index de texture.

5.3.2 Résultats

Pour vérifier si la texture mesurée à différentes échelles est cohérente, dans un premier temps, nous avons comparé les index issus des images TEP et ceux issus des images autoradiographiques. Les résultats sont résumés dans le Tableau 5.3. Tous les index sont significativement corrélés entre les images TEP et AR, ARR ou ARRL, et cette corrélation augmente entre la comparaison TEP vs. AR ($r_s=[0,57-0,85]$) et TEP vs. ARRL ($r_s=[0,77-0,95]$). Néanmoins, les résultats des tests de Wilcoxon montrent que les valeurs absolues des index de texture diffèrent entre les deux modalités, sauf pour LGZE et HGZE (après la correction de Holm) dans TEP vs. ARR. Comme le montre la Figure 5.3, les images AR sont plus homogènes que les images TEP (Homogeneity plus élevé, Entropy et HGZE plus faibles) alors que les images ARR et ARRL semblent plus hétérogènes (Homogeneity plus faible, Entropy et HGZE plus élevés).

Index	TEP vs. AR		TEP vs. ARR		TEP vs. ARRL	
	r_s	pvalue	r_s	pvalue	r_s	pvalue
Homogeneity	0,66†	2,62e-16*	0,69†	3,42e-07*	0,84†	2,71e-05*
Entropy	0,57†	6,60e-07*	0,91†	0,0035*	0,95†	0,0057*
SRE	0,67†	2,62e-16*	0,67†	3,49e-05*	0,77†	4,47e-05*
LRE	0,70†	2,62e-16*	0,64†	2,44e-07*	0,89†	0,0002*
LGZE	0,83†	0,0009*	0,76†	9,9935	0,84†	0,0002*
HGZE	0,85†	0,0171*	0,70†	0,0459	0,89†	0,0003*
Max	0,75†	2,62e-16*	0,72†	7,10e-13*	0,83†	1,03e-09*

Tableau 5.3 – Coefficients de corrélation de Spearman (r_s) et p-values du test de Wilcoxon pour la comparaison entre la texture mesurée sur les images TEP et celle mesurée sur les images autoradiographiques. Les p-values en gras sont inférieures à 5%. † signifie que la corrélation est significative au seuil de 5%. Les p-values en gras sont significatives au seuil de 5%. * signifie que la p-value reste significative après correction de Holm.

Dans un second temps, nous avons comparé la texture mesurée sur les coupes après marquage à l'hémaroxyline (H et HL) avec la texture issue des images TEP ou AR (cf. Tableau 5.4). Entre H et AR, aucun index de texture ni Max n'est significativement corrélé positivement et leurs valeurs absolues sont différentes. Quand on lisse les images H, tous les index, y compris Max sont significativement corrélés ($r_s=[0,53-0,67]$) mais seules les valeurs absolues de Entropy et Max ne diffèrent pas significativement. Les images AR semblent plus hétérogènes que les images HL comme l'illustre la Figure 5.4.

La même analyse a été réalisée entre la texture issue de H ou HL et celle mesurée sur les images TEP (cf. Tableau 5.4). Aucun index n'est significativement corrélé entre les images TEP et H et seulement SRE, LGZE, HGZE et Max sont significativement corrélés entre TEP et HL. On note également que les valeurs absolues de tous les index diffèrent entre les deux modalités. Les images HL semblent plus homogènes que les images TEP (cf. Figure 5.4).

Index	H vs. AR		HL vs. AR		H vs. PET		HL vs. PET	
	r_s	pvalue	r_s	pvalue	r_s	pvalue	r_s	pvalue
Homogeneity	-0,18	2,62e-16*	0,57†	1,24e-06*	-0,23	0,0141*	0,35	2,62e-16*
Entropy	0,31	2,62e-16*	0,55†	0,5858	0,13	2,54e-14*	0,27	3,82e-07*
SRE	-0,25	2,62e-16*	0,55†	0,0148*	-0,13	1,94e-07*	0,39†	2,62e-16*
LRE	-0,41†	2,62e-16*	0,56†	2,54e-14*	-0,54†	2,28e-06*	0,35	2,62e-16*
LGZE	-0,23	0,0054*	0,67†	0,0001*	-0,06	7,49e-08*	0,58†	1,23e-08*
HGZE	0,32	7,22e-05*	0,53†	0,0026*	0,29	0,0726	0,39†	8,45e-05*
Max	0,20	2,62e-16*	0,66†	0,6665	0,09	2,62e-16*	0,38†	2,62e-16*

Tableau 5.4 – Coefficients de corrélation de Spearman (r_s) et p-values du test de Wilcoxon pour la comparaison entre la texture mesurée sur les images TEP ou sur les images autoradiographiques et celle mesurée sur les coupes après marquage à l'hémaroxyline. † signifie que la corrélation est significative au seuil de 5%. Les p-values en gras sont significatives au seuil de 5%. * signifie que la p-value reste significative après correction de Holm.

5.3.3 Discussion

Entre les images TEP et les images autoradiographiques, nous observons que tous les index sont plus corrélés entre TEP et ARRL que entre TEP et AR. Ces résultats sont cohérents avec les conclusions obtenues précédemment puisque les images ARRL sont ré-échantillonnées et lissées pour correspondre à la taille des voxels et à la résolution spatiale des images TEP. Néanmoins, la corrélation entre la texture issue des images TEP et des images AR est significative pour tous les paramètres, démontrant pour la première fois la capacité des index à quantifier une hétérogénéité vue à deux échelles différentes. Pour Entropy, nous retrouvons bien que l'index est fortement influencé par la différence de taille de voxel (TEP vs. AR, $r_s = 0,57$) et très peu par la différence en terme de résolution spatiale (TEP vs. ARRL, $r_s = 0,91$).

Même si nous avons observé une texture similaire entre les images TEP et autoradiographiques, nous avons montré que les index issus de ces deux types d'images ne reflètent pas précisément l'hétérogénéité de distribution des cellules tumorales sur les images après marquage à l'hémaroxyline (H).

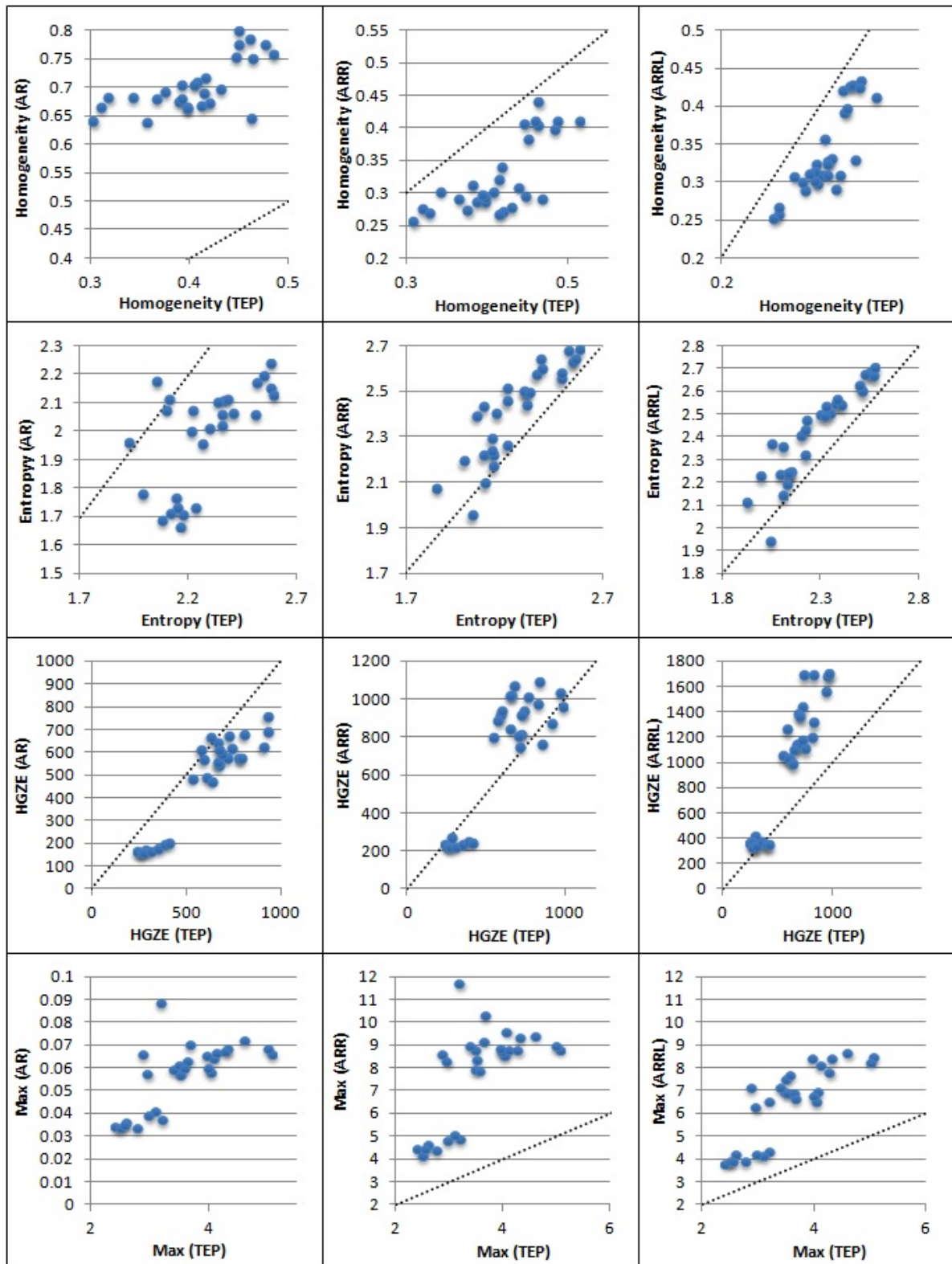


FIGURE 5.3 – Graphiques des index issus des images TEP et des images autoradiographiques. Les lignes en pointillés correspondent à la première bissectrice.

En effet, ce n'est pas la répartition précise des cellules qui est quantifiée par l'analyse de texture sur les images TEP ou AR, mais une hétérogénéité plus grossière. Cette différence peut provenir de la

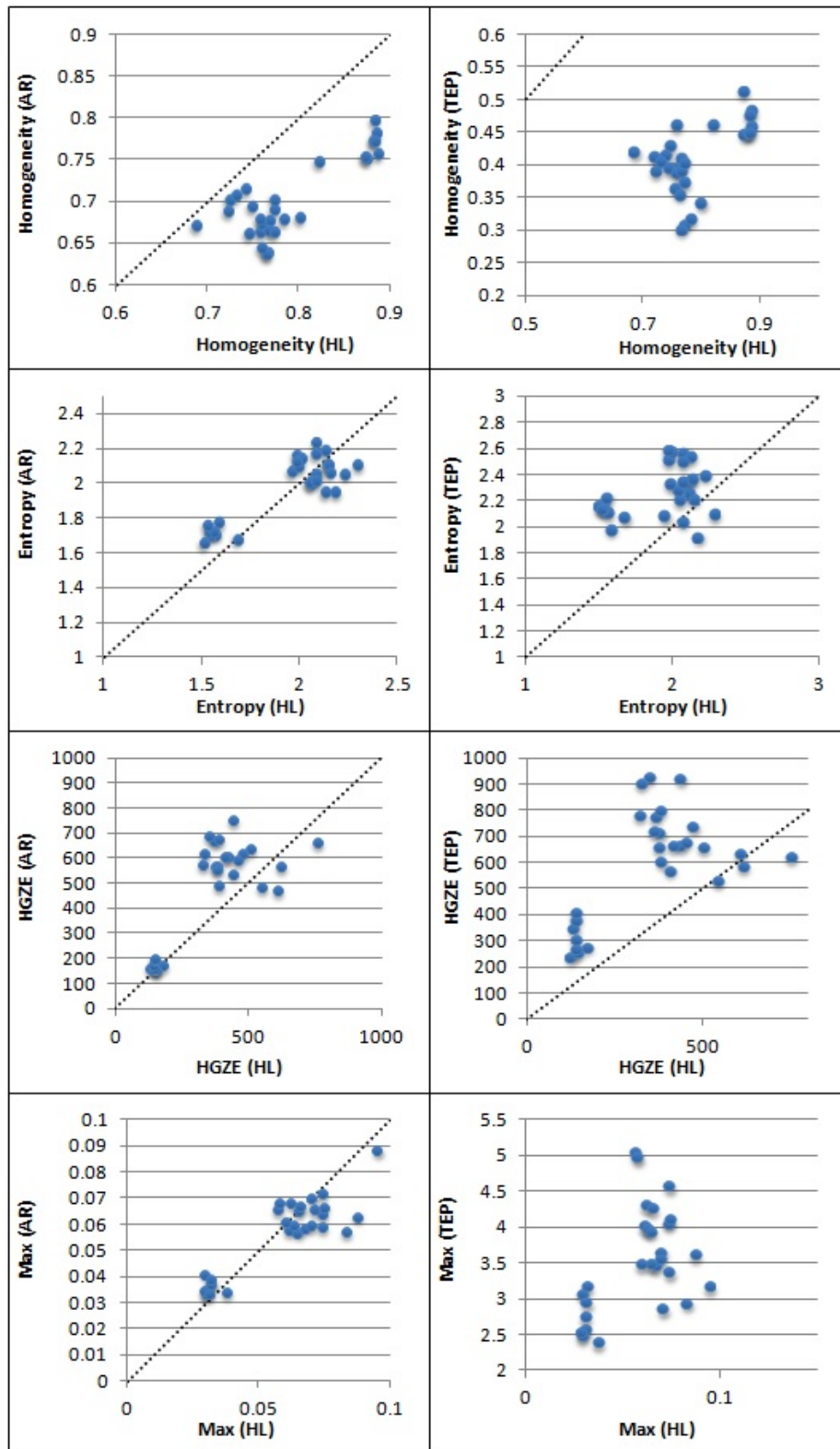


FIGURE 5.4 – Graphiques des index issus des images TEP ou autoradiographiques et des coupes histologiques. Les lignes en pointillés correspondent à la première bissectrice.

différence de résolution spatiale puisque la corrélation augmente entre les index issus des images TEP ou AR et ceux mesurés sur les images H lissées (HL), mais aussi de la non-spécificité de fixation du

18F-FDG par rapport au marquage à l'hématoxyline. En effet, à partir des images TEP au 18F-FDG ou sur les images autoradiographiques, l'hétérogénéité quantifiée correspond à l'hétérogénéité de la sur-expression des protéines membranaires assurant le transport du glucose (GLUT) et des enzymes permettant d'initier la glycolyse (hexokinase), alors que les images histochimiques rendent compte de la répartition des noyaux. Même si la mesure de l'hétérogénéité est imparfaite, l'analyse de texture reste toutefois informative et permet de caractériser la structure du tissu tumoral.

Dans une étude récente, [Rajkumar et al., 2015] montrent que l'analyse de texture des images de tomographie d'émission monophotonique (TEMP) permet de différencier des types de cancers métastatiques colorectaux avec des valeurs pour Homogeneity plus élevées pour le type bien différencié par rapport au type peu différencié. Les auteurs ont montré que, parmi les métastases hépatiques peu différenciées, l'analyse de texture permettait de quantifier l'action d'un traitement anti-angiogénique avec, par exemple, une augmentation de l'index Homogeneity. À l'inverse, les index de texture ne permettent pas de différencier les lésions traitées et non-traitées parmi les métastases bien différenciées. Cet exemple illustre la quantification imparfaite de l'hétérogénéité cellulaire dans les tumeurs à partir des images TEMP puisque l'analyse histologique montre des modifications induites par le traitement dans les deux types de cancers alors que l'analyse de texture ne quantifie ces modifications que pour les lésions peu différenciées. Avant un passage en clinique de l'analyse de texture, il va donc être indispensable de comprendre comment la structure des tissus et l'architecture cellulaire observables au microscope influencent la distribution des niveaux de gris sur les images TEMP ou TEP. Des analyses complémentaires doivent permettre d'expliquer le lien entre l'hétérogénéité biologique et l'hétérogénéité de fixation. Il est encore nécessaire de clarifier le potentiel de l'analyse de texture pour capturer des informations sur le phénotype tumoral et fournir des biomarqueurs pour prédire ou quantifier la réponse aux traitements.

Dans une autre étude, [Hoeben et al., 2014] se sont intéressés à la capacité des index issus de l'histogramme à distinguer différentes lignées tumorales dans des cancers ORL. Ils ont montré que les index issus des images TEP au 18F-FDG n'étaient pas capables de différencier les lignées tumorales, contrairement aux mêmes index mesurés directement sur les coupes après des marquages immunohistochimiques. L'utilisation des index de texture pourrait permettre de faire cette distinction puisque [Rajkumar et al., 2015] rapportent une texture différente entre deux sous-types de cancers colorectaux, les index issus de l'histogramme n'étant pas sensibles à l'organisation spatiale des niveaux de gris à l'intérieur du volume tumoral.

5.4 Impact de la densité cellulaire sur l'analyse de texture

Même si la texture mesurée sur les images TEP ou autoradiographiques n'est pas le reflet exact de la texture issue des coupes après marquage à l'hématoxyline, nous avons cherché à déterminer si les index de texture étaient sensibles localement à des différences de densité cellulaire.

5.4.1 Matériels & méthodes

Nous avons choisi une coupe présentant une grande variabilité locale de densité cellulaire afin d'en étudier l'impact sur l'analyse de texture. À partir de l'image H de cette coupe (voxel = $50 \times 50 \times 20 \mu m^3$), nous avons effectué un seuillage afin de binariser l'image pour identifier la présence de cellules. Nous avons alors tracé des sous-régions de 30×30 voxels et calculé le pourcentage de cellules dans ces régions (cf. Figure 5.5). Dans l'image AR correspondante, nous avons calculé les six index de texture (Homogeneity, Entropy, SRE, LRE, LGZE et HGZE) après une discrétisation absolue avec 64 niveaux dans ces mêmes sous-régions. Nous avons alors étudié la capacité des index de texture à

distinguer différents pourcentages de cellules regroupés en classe à l'aide de boîtes à moustache et d'un test de Kruskal-Wallis avec une correction de Holm pour les comparaisons multiples.

Pour les index sensibles aux micro-hétérogénéités (Homogeneity, Entropy, SRE et LRE) que nous avons identifiés à la section 3.5 et pour lesquels nous avons montré qu'ils n'étaient pas sensibles au type de fixation (hypofixation ou hyperfixation) des hétérogénéités simulées, nous avons séparé les sous-régions en trois classes, choisissant les pourcentages seuils apparaissant ci-dessous de sorte d'équilibrer le nombre de sous-régions par classe :

- les sous-régions très hétérogènes qui correspondent à des zones composées à la fois de cellules et de matrice extracellulaire (pourcentage de cellules entre 29 et 64%, $n=28$ sous-régions) ;
- les sous-régions plutôt homogènes qui sont composées soit majoritairement de cellules, soit majoritairement de matrice extracellulaire, dont le pourcentage de cellules est respectivement compris entre 64 et 85% ($n=14$) ou entre 11 et 29% ($n=14$) ;
- les sous-régions très homogènes qui contiennent soit presque exclusivement des cellules (pourcentage de cellules entre 85 et 100%, $n=15$), soit presque exclusivement de la matrice extracellulaire (pourcentage de cellules entre 0 et 11%, $n=15$).

Pour les index sensibles aux macro-hétérogénéités (LGZE et HGZE, cf. section 3.5) ainsi que Max, et qui permettent la distinction entre des hétérogénéités hyperfixantes ou hypofixantes dans les simulations, nous avons créé trois nouvelles classes avec une augmentation progressive du pourcentage de cellules dans la sous-région :

- les sous-régions composées principalement de matrice extracellulaire (pourcentage de cellules entre 0 et 29%, $n=28$ sous-régions) ;
- les sous-régions mixtes comportant à la fois des cellules et de la matrice extracellulaire (pourcentage de cellules entre 29 et 64%, $n=28$) ;
- les sous-régions composées principalement de cellules (pourcentage de cellules entre 64 et 100%, $n=30$).

5.4.2 Résultats

Sur une coupe sélectionnée pour sa grande diversité de pourcentage cellulaire, nous avons calculé les index de texture sur l'image AR dans des sous-régions.

Pour Homogeneity, Entropy, SRE et LRE, nous avons tracé la valeur de l'index en fonction de trois types de sous-régions : très hétérogènes, plutôt homogènes ou très homogènes, comme présenté sur la Figure 5.6. Nous observons que les index Homogeneity et LRE sont plus élevés dans les sous-régions très homogènes qui contiennent quasi-exclusivement soit des cellules soit de la matrice extracellulaire, par rapport aux sous-régions très hétérogènes qui comportent environ 50% de cellules. Le comportement strictement opposé est observé pour Entropy et SRE. Pour tous les index, les tests de Kruskal-Wallis sont significatifs ($p\text{-value}<5\%$) pour la séparation des trois types de sous-régions.

Pour LGZE, HGZE et Max, nous avons regroupé les pourcentages en trois autres types de sous-régions : les sous-régions majoritaires en matrice extracellulaire, mixtes ou majoritaires en cellules. La Figure 5.7 montre que LGZE diminue et HGZE et Max augmentent quand on augmente le pourcentage de cellules dans la sous-région, avec des tests de Kruskal-Wallis significatifs pour les trois index.

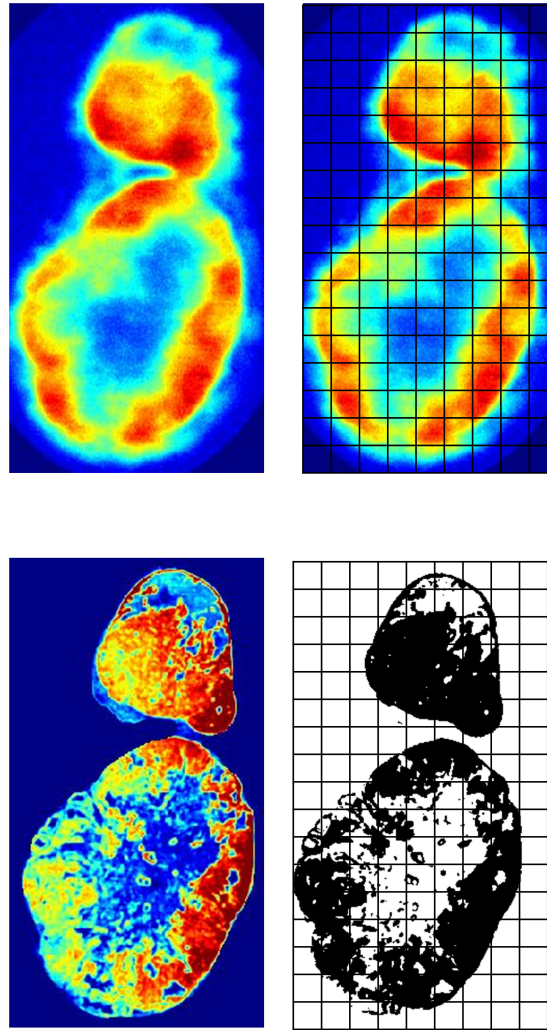


FIGURE 5.5 – Schéma de découpage des sous-régions de 30×30 voxels sur une image AR (en haut) et sur la coupe correspondante après marquage à l'hématoxyline et seuillage (en bas).

5.4.3 Discussion

Cette analyse nous a permis de montrer que les index de texture mesurés sur les images AR étaient influencés par le pourcentage de cellules présentes dans la région d'intérêt. Les paramètres se séparent en deux types : ceux qui sont sensibles à l'augmentation progressive du pourcentage de cellules comme HGZE, LGZE et Max, et ceux qui sont plutôt sensibles à une répartition mixte (50% de cellules, 50% de matrice extracellulaire) par rapport à une répartition déséquilibrée (dominante ou minorité de cellules) comme Homogeneity, Entropy, SRE et LRE.

L'identification de ces deux groupes d'index est cohérente avec ceux établis à partir des sphères simulées (cf. section 3.5). Nous avons montré que HGZE et LGZE étaient sensibles aux macro-hétérogénéités et capables de distinguer l'ajout d'hyperfixations ou d'hypofixations, comme ici où ils discernent les régions minoritaires des régions majoritaires en cellules. À l'inverse, Homogeneity, Entropy, SRE et LRE étaient plutôt sensibles aux micro-hétérogénéités et ne permettaient pas de distinguer l'ajout d'hyperfixations et d'hypofixations. Ce constat se retrouve ici où les index sont capables de séparer des régions mixtes, de régions déséquilibrées sans distinguer la « nature » du déséquilibre (majorité ou minorité de cellules).

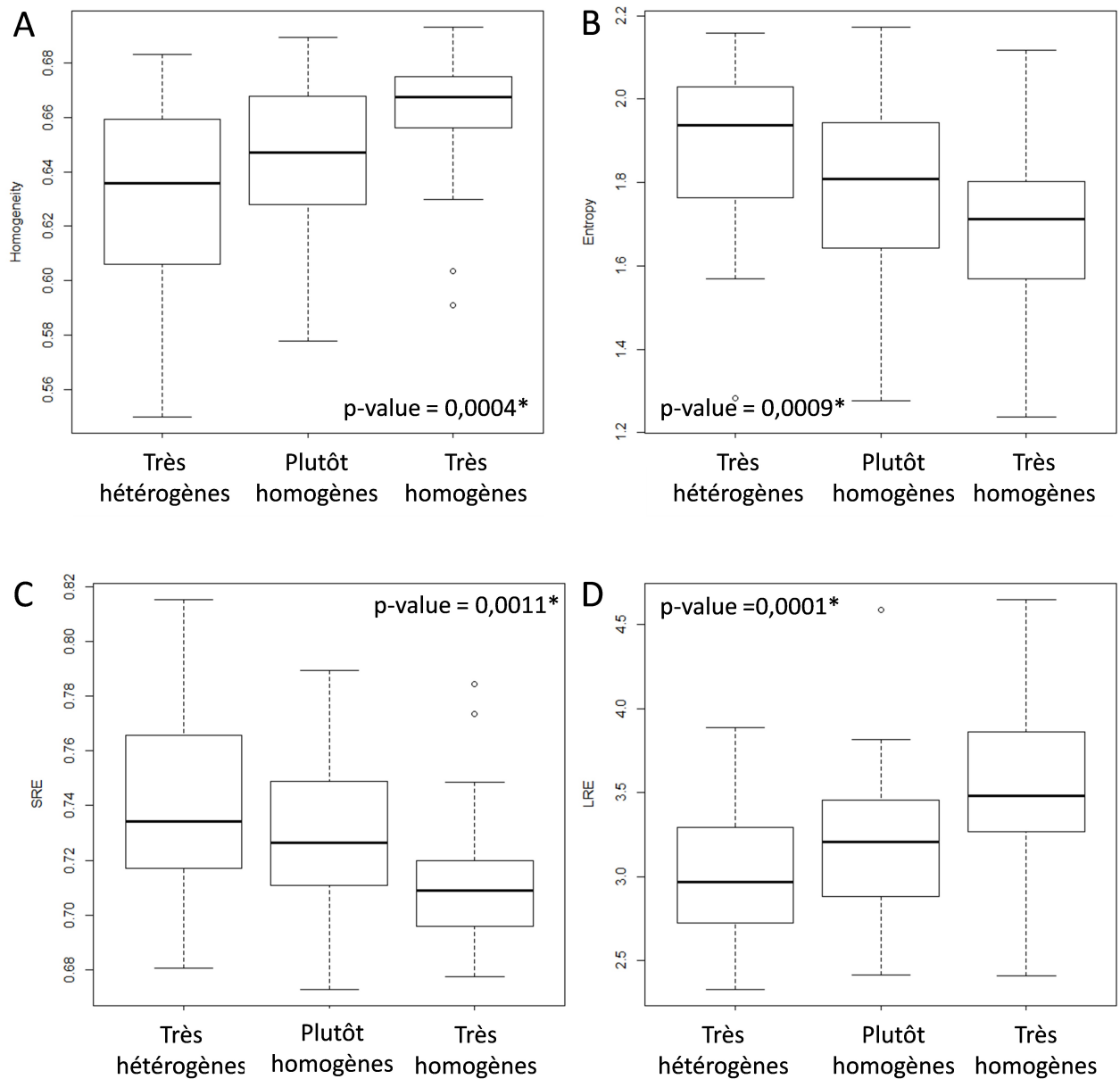


FIGURE 5.6 – Boîtes à moustache et p-values du test de Kruskal-Wallis pour les index Homogeneity, Entropy, SRE et LRE suivant 3 types de sous-régions : très hétérogènes (pourcentage de cellules entre 29 et 64%), plutôt homogènes (soit entre 11 et 29%, soit entre 64 et 85% de cellules) ou très homogènes (soit entre 0 et 11%, soit entre 85 et 100% de cellules). * signifie que la p-value reste significative après correction de Holm.

Dans cette étude, les paramètres HGZE et LGZE présentent de meilleures performances que Homogeneity, Entropy, SRE et LRE pour séparer les sous-régions. Ils conduisent également à des p-values plus faibles que celle de Max. Ce résultat montre que ces deux index sont à la fois sensibles au niveau de gris dans l'image mais également à leur répartition, à la différence de Max qui est une mesure ponctuelle sur un unique voxel et qui ne prend pas en compte l'hétérogénéité de la distribution sur les 900 voxels (30×30) de la sous-région.

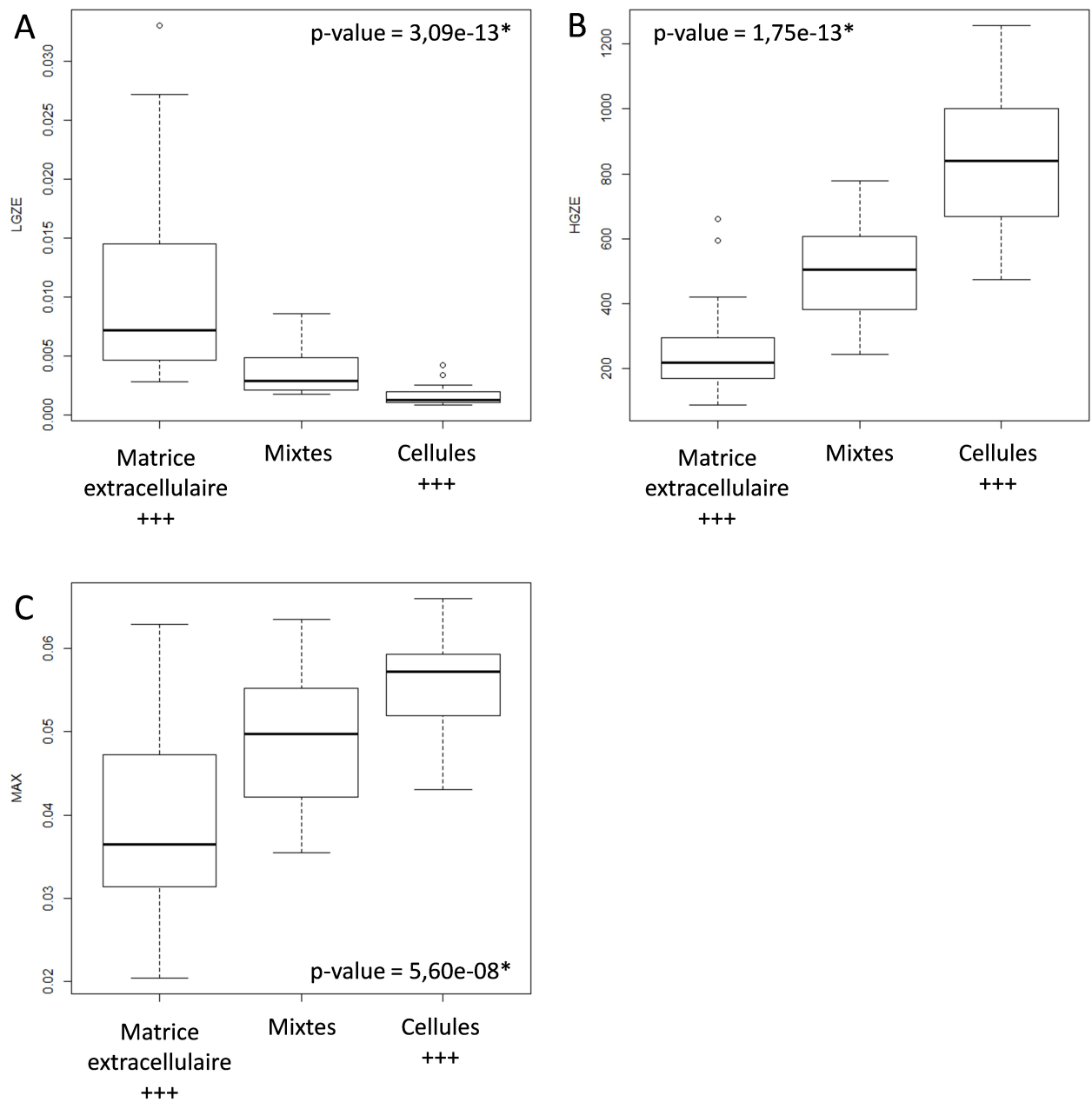


FIGURE 5.7 – Boîtes à moustache et p-values du test de Kruskal-Wallis pour les index LGZE, HGZE et Max suivant 3 types de sous-régions : majoritaires en matrice extracellulaire (pourcentage de cellules entre 0 et 29%), mixtes (entre 29 et 64% de cellules), majoritaires en cellules (entre 64 et 100% de cellules). * signifie que la p-value reste significative après correction de Holm.

Ce chapitre nous a permis d'étudier à la fois l'impact de la taille des voxels et de la résolution spatiale sur les index, mais également de montrer une forte cohérence entre l'analyse de texture réalisée *in vivo* sur les images TEP et *ex vivo* sur les images autoradiographiques. Dans un second temps, grâce à l'étude des coupes histologiques, nous avons montré que, même si les index de texture ne capturent pas point-par-point l'organisation fine de la tumeur du fait de la résolution spatiale limitée et du manque de spécificité du traceur, les informations extraites quantifient grossièrement la structure tumorale et sont sensibles à la distribution des cellules. Ces paramètres de texture peuvent donc potentiellement être utiles pour évaluer l'agressivité tumorale ou les résistances aux traitements. Les caractéristiques tumorales issues de l'analyse de texture permettent d'enrichir le profil tumoral en complément des informations fournies par les biopsies. Toutefois, l'emploi de ces index requiert des analyses complémentaires pour comprendre davantage leur lien avec l'organisation cellulaire et être utilisé judicieusement en clinique.

Discussion générale

Comme nous l'avons mentionné au Chapitre 2, l'analyse de texture des images TEP en oncologie connaît un intérêt croissant depuis 2009. Le calcul de biomarqueurs directement à partir des données d'imagerie pourrait permettre d'évaluer la réponse aux traitements ou de prédire la survie du patient afin d'adapter sa prise en charge. Cette recherche s'inscrit dans le domaine de la radiomique qui vise à extraire et combiner des informations, non quantifiables par une analyse visuelle, pour enrichir le profil de la tumeur. Dans ce concept, la sélection des paramètres d'intérêt se fait par rapport à leur pertinence et leur robustesse.

Dans le cas de l'analyse de texture des images TEP, nous avons donc d'abord étudié l'impact des différents facteurs intervenant lors du calcul des index de texture afin d'optimiser cette première étape.

Comment calculer au mieux les paramètres de texture ? Lesquels choisir ?

Dans un premier temps, la construction des groupes de corrélation (cf. Chapitre 3) nous a permis de réduire le nombre de paramètres à calculer, limitant l'apparition de résultats faux positifs [Chalkidou et al., 2015], et d'expliquer les résultats publiés en analyse de texture. Par exemple, nous avons montré que si l'index GLNUz est présenté comme prédictif de la réponse à la thérapie pour les cancers de l'œsophage dans [Tixier et al., 2011], c'est en fait parce qu'il est très fortement corrélé avec VM (cf. section 3.2), qui lui même a été identifié par les mêmes investigateurs comme un facteur fortement prédictif de la réponse chez des patients similaires [Hatt et al., 2011]. Ces investigateurs ont donc conclu à tort que l'index GLNUz reflétait une texture tumorale prédictive de la réponse à la thérapie, alors que c'est le volume de la tumeur qui l'est. L'identification des groupes de corrélation permet donc d'analyser plus finement les conclusions rapportées.

Dans un second temps, la compréhension du lien entre la corrélation des index avec VM et la méthode de discrétisation, nous a permis de proposer une méthode alternative, la discrétisation absolue, et ainsi de rendre plus pertinente l'analyse de texture, y compris pour de petits volumes tumoraux, à la différence de ce qui avait été rapporté dans [Brooks and Grigsby, 2014] et [Brooks and Grigsby, 2015].

À l'issue du Chapitre 3, nous avons établi un ensemble de recommandations concernant le calcul des index de texture :

- 1°) utiliser une méthode de segmentation automatique ou semi-automatique ;
- 2°) discrétiser les intensités avec 64 niveaux de gris en utilisant une discrétisation absolue ;
- 3°) choisir un index par groupe ;
- 4°) toujours comparer les résultats avec ceux des index conventionnels.

Nous avons sélectionné six index de texture, représentatifs des six groupes de corrélation et choisi pour leur robustesse vis-à-vis de la méthode de segmentation : Homogeneity, Entropy, SRE, LRE, LGZE et HGZE, en complément des cinq index conventionnels (SUVmax, SUVmean, SUVpeak, VM et TLG).

À partir de ces index de texture sélectionnés, nous avons cherché lesquels étaient les plus prometteurs pour une utilisation en clinique.

Quelles sont les performances des index sélectionnés ?

Dans les Chapitres 4 et 5, nous avons étudié les corrélations entre les caractéristiques tumorales et ces six index de texture et nous avons comparé la texture mesurée à différentes échelles. Les résultats obtenus sont résumés dans le Tableau 5.5, après une discrétisation absolue entre 0 et 20 unités SUV sur 64 niveaux. En parallèle, le Tableau 5.6 récapitule les résultats présentés dans la littérature pour ces 6 index de texture.

Nous allons analyser ces conclusions en étudiant séparément les index identifiés comme sensibles aux macro-hétérogénéités et ceux sensibles aux micro-hétérogénéités, comme présenté à la section 3.5.

• Les index de macro-hétérogénéités : LGZE et HGZE

Les index LGZE et HGZE présentent les mêmes tendances sur l'ensemble des tests effectués pendant ce travail de thèse (cf. Tableau 5.5). Ces résultats sont cohérents avec ceux présentés au Chapitre 3 où les simulations ont mis en évidence un comportement similaire (sensibles aux macro-hétérogénéités) mais en sens opposé pour ces deux index. Toutefois, avec 8 tests sur 18 inférieurs au seuil de 5% avec la correction de Holm et 3 tests significatifs sans correction, aucune p-value n'est plus faible que celles obtenues par les index conventionnels : mesures SUV, VM ou TLG. La forte corrélation entre LGZE ou HGZE et SUVmax avec la discrétisation absolue peut expliquer ces résultats ($|r_s|=0,99$ dans VOI-T contre $|r_s|=[0,89-0,95]$ dans VOI-F pour la cohorte CPNPC-2, cf. Tableau 3.8). Toutefois, au Chapitre 5, LGZE et HGZE évoluent proportionnellement au pourcentage de cellules avec de meilleures performances que l'intensité maximale mesurée sur des coupes autodiagraphiques. Dans ce même Chapitre, nous avons mis en évidence que LGZE était moins influencé que HGZE à la fois par la différence de taille de voxels et de résolution spatiale entre deux images. Ces résultats sont cohérents avec ceux rapportés très récemment par [Yan et al., 2015] et résumés dans le Tableau 5.7. Les auteurs montrent que LGZE est plus robuste aux différents paramètres de reconstruction que HGZE.

Dans la littérature, ces deux paramètres ont été très peu étudiés. [Lapa et al., 2015] ont rapporté chez 12 patients que HGZE permettait de prédire la progression pour les lésions thyroïdiennes avec de meilleures performances que SUVmax ou SUVmean, VM n'étant pas testé (cf. Tableau 5.6). Le comportement inverse est observé dans [Tixier et al., 2014b] qui montrent, dans des cancers pulmonaires, que HGZE n'est pas prédictif de la survie contrairement aux mesures SUV, VM et TLG. Chez 20 patientes porteuses de cancers du col de l'utérus, [Yang et al., 2013] ont montré que l'index HGZE changeait significativement au cours du traitement dans le groupe des patientes répondeuses à la thérapie, alors que LGZE était significativement différent dans le groupe des patientes, soit partiellement répondeuses, soit en progression. Les mesures SUV permettent quant à elles de suivre l'évolution au cours du traitement dans les deux groupes de patients. Dans [Tixier et al., 2012, Leijenaar et al., 2013], LGZE apparaît moins robuste que HGZE aux tests-retests.

Des études complémentaires semblent donc nécessaires pour valider l'utilisation ou non de ces deux paramètres en réalisant des analyses conjointes avec les mesures SUV, afin de déterminer leur valeur ajoutée. On peut considérer qu'actuellement, le niveau de preuve concernant ces deux index est insuffisant pour conclure à leur valeur ajoutée dans les cancers thyroïdiens ou du col de l'utérus. Cependant, ils méritent très certainement d'être étudiés plus systématiquement et de ne pas être seule-

ment rapportés quand ils présentent de bonnes performances.

• **Les index de micro-hétérogénéités : Homogeneity, Entropy, SRE et LRE**

L'analyse du Tableau 5.5 permet de mettre en évidence que, parmi les 4 index identifiés comme sensibles aux micro-hétérogénéités au Chapitre 3, Entropy conduit aux meilleures performances. En effet, sur 18 tests réalisés, Entropy conduit à une p-value non significative dans seulement deux tests, où, de plus, aucun autre index n'est significatif (distinction des lésions parmi les gliomes de bas grade et identification de la surexpression de HER-2). Cet index conduit même à des performances de discrimination qui semblent être meilleures que celles des index conventionnels dans 6 cas (p-values inférieures sans que la différence avec les index conventionnels ne soit statistiquement significative). Les résultats des Chapitres 3 et 5 ont également mis en évidence un impact modéré de la différence de taille de voxels ainsi qu'une très bonne robustesse vis-à-vis de la résolution spatiale ou de la méthode de segmentation. De surcroît, cet index est sensible à l'hétérogénéité biologique (cf. Chapitre 5).

Les bonnes performances d'Entropy, sa définition attractive de mesure du désordre et sa grande robustesse font que cet index est le plus calculé dans la littérature parmi les six index que nous avons identifiés (cf. Tableau 5.5). Ainsi, Entropy est présenté comme un facteur pronostique pour les cancers pulmonaires dans [Pyka et al., 2015], [Hatt et al., 2015] ou [Tixier et al., 2014b], thyroïdiens [Lapa et al., 2015] ou ORL [Cheng et al., 2013], mais pas dans les cancers de l'œsophage comme rapporté dans [Hatt et al., 2015]. Cet index permettrait également d'identifier les lésions répondant à la thérapie dans des cancers de l'œsophage d'après [Tixier et al., 2011], et dans les cancers du col de l'utérus (n=14 patientes) [El Naqa et al., 2009] mais pas dans le cas des cancers ORL (n=9 patients) [El Naqa et al., 2009]. Entropy est significativement corrélé au stade tumoral dans les cancers de l'œsophage [Dong et al., 2013] et dans les cancers pulmonaires [van Gómez López et al., 2014] et permet aussi la distinction entre les tumeurs bénignes et malignes dans les lésions des os et des tissus mous [Xu et al., 2014]. Entropy conduit même à de meilleures performances que les index conventionnels dans 5 études sur 12 d'après le Tableau 5.5. Finalement, Entropy est modérément influencé par les paramètres d'acquisition et de reconstruction comme le démontrent [Galavis et al., 2010] puis [Yan et al., 2015] (cf. Tableau 5.7), très stable lors des tests-retests d'après [Tixier et al., 2012] ou [Leijenaar et al., 2013] et peu influencé par la correction de l'effet de volume partiel [Hatt et al., 2013].

Toutefois, ces publications, quand cela est spécifié, utilisent une discrétisation relative (excepté dans [Leijenaar et al., 2013]). Or, comme nous l'avons démontré au Chapitre 3 et comme cela avait été rapporté dans [Brooks and Grigsby, 2014], l'index Entropy, calculé après une discrétisation relative, est fortement corrélé au VM, en particulier pour les petits VOI. Ces conclusions nécessitent donc d'être confirmées avec l'utilisation de la discrétisation absolue. Néanmoins, au vu de nos résultats et de ceux de la littérature, cet index semble très prometteur pour la caractérisation de l'hétérogénéité tumorale à partir des images TEP.

Concernant les autres index sensibles aux micro-hétérogénéités, SRE et LRE conduisent à des performances quasiment similaires sur les 18 tests effectués (cf. Tableau 5.5), alors que Homogeneity conduit à de meilleures capacités de distinction avec notamment trois p-values inférieures à celles des index conventionnels. Dans les Chapitres 3 et 5, nous avons montré que ces trois index étaient modérément affectés par les différences entre deux images en termes de taille de voxels et de résolution spatiale, et robustes vis-à-vis de la méthode de segmentation choisie. Ils sont également sensibles à l'hétérogénéité biologique, comme nous l'avons vu au Chapitre 5.

Dans la littérature, parmi Homogeneity, SRE et LRE, Homogeneity est celui qui est le plus souvent rapporté, excepté dans [Lapa et al., 2015] qui montre que SRE est un facteur pronostique dans les

cancers thyroïdiens chez 12 patients avec de meilleures performances que les mesures SUV. Homogeneity semble prédictif de la réponse à la thérapie dans les cancers de l'œsophage [Tixier et al., 2011] (AUC=0,89), ORL [El Naqa et al., 2009] (AUC=0,83) et, dans une moindre mesure, dans les cancers du col de l'utérus [El Naqa et al., 2009] (AUC=0,62). Homogeneity serait également un facteur pronostique dans les cancers pulmonaires [Tixier et al., 2014b]. Toutefois Homogeneity, SRE et LRE sont sensibles aux paramètres d'acquisition et de reconstruction des images TEP d'après [Galavis et al., 2010], ce qui a été confirmé par [Yan et al., 2015] (cf. Tableau 5.7), mais très robustes lors des études tests-retests [Leijenaar et al., 2013]. Parmi ces trois index, sur la base de la littérature et de nos résultats, Homogeneity semble le plus prometteur mais le manque de données dans la littérature sur SRE et LRE ne nous permet pas de les exclure à l'heure actuelle.

En comparant les deux types d'index, ceux qui sont sensibles aux micro-hétérogénéités, comme Entropy et Homogeneity, semblent plus performants que les index sensibles aux macro-hétérogénéités, comme HGZE et LGZE, à la fois d'après nos résultats et ceux de la littérature, même si peu d'études rapportent systématiquement les six index que nous avons sélectionnés. Toutefois, il apparaît, dans les deux catégories d'index, des exemples où les paramètres conduisent à de meilleures performances que les index conventionnels comme les mesures SUV ou VM. Ceci suggère que l'analyse de texture des images TEP pourrait quantifier une information pertinente pour la caractérisation de l'hétérogénéité tumorale ainsi que prédire la réponse aux traitements et la survie du patient. Avant la réalisation d'études prospectives, il va être essentiel de s'intéresser aux valeurs absolues de ces paramètres qui sont fortement influencées par le protocole de calcul des matrices de texture. À l'heure actuelle, ces valeurs absolues ne sont pas systématiquement reportées dans les publications ; certains auteurs normalisent les valeurs entre 0 et 1 sur leur cohorte [Tixier et al., 2014b] ; d'autres ne reportent que la médiane [Yang et al., 2013]. L'établissement de seuils au-delà ou en deça desquels la tumeur pourra être considérée comme hétérogène sera pourtant indispensable avant un passage en clinique.

			Homogeneity	Entropy	SRE	LRE	LGZE	HGZE	SUVmax	SUVmean	SUVpeak	VM	TLG
Robustesse	Cancers pulmonaires non à petites cellules	Méthode de segmentation											
	Autoradiographie	Taille des voxels					*			-	-	-	-
		Résolution spatiale	*	*						-	-	-	-
Applications cliniques	Cancers pulmonaires non à petites cellules et cancers mammaires	Tissu tumoral vs. Tissu hépatique	*	*	*	*						-	
		Tissu tumoral vs. Tissu controlatéral										-	
		Tissu hépatique vs. Tissu controlatéral										-	
		Tumeurs pulmonaires vs. Tumeurs mammaires											
	Cancers mammaires	Corrélation avec l'analyse visuelle de l'hétérogénéité tumorale											
		Récepteurs aux œstrogènes											
		Récepteurs à la progestérone											
		Surexpression de HER-2											
		Lésions triple-négatives vs. Non triple-négatives											
	Cancers pulmonaires non à petites cellules	Stade tumoral											
		Stade ganglionnaire											
		Sous-types de cancer : épidermoïdes vs. adénocarcinomes	*	*									
		Niveau de différenciation		*									
		Présence d'une nécrose histologique		*									
	Gliomes	Lésions nouvellement diagnostiquées vs. en récurrence parmi les tumeurs de haut grade		*									
		Lésions nouvellement diagnostiquées vs. en récurrence parmi les tumeurs de bas grade											
		Haut grade vs. bas grade pour les lésions nouvellement diagnostiquées	*	*	*	*							
		Haut grade vs. bas grade pour les lésions en récurrence											

Tableau 5.5 – Tableau résumant les conclusions obtenues. Vert : p-value inférieure à 5% avec la correction de Holm ou coefficient de corrélation $r=[0,80-1]$. Jaune : p-value inférieure à 5% sans la correction de Holm ou $r=[0,60-0,80[$. Rouge : p-value supérieure à 5% ou $r < 0,60$. * signifie que p-value (index de texture) < p-value (index conventionnels) ou r (index de texture) > r (index conventionnels).

			Homogeneity	Entropy	SRE	LRE	LGZE	HGZE	SUVmax	SUVmean	SUVpeak	VM	TLG
Différenciation bénin vs. malin	[Xu et al. 2014]	Lésions des os et des tissus mous	-	*	-	-	-	-			-	-	-
Corrélation avec le stade tumoral	[van Gomez Lopez et al. 2014]	Cancers pulmonaires non à petites cellules			-	-	-	-			-		
	[Dong et al. 2013]	Cancers de l'œsophage	-	*	-	-	-	-		-	-	-	-
Prédiction de la réponse à la thérapie	[Tixier et al. 2011] [Hatt et al. 2013]	Cancers de l'œsophage			-	-	-				-		-
	[El Naqa et al. 2009]	Cancers du col de l'utérus	*	*	-	-	-	-			-		-
	[El Naqa et al. 2009]	Cancers ORL			-	-	-	-			-		-
Évolution de l'hétérogénéité	[Yang et al. 2013]	Cancers du col de l'utérus	-	-							-	-	-
Prédiction de la survie	[Cheng et al. 2013]	Cancers ORL	-		-	-	-	-		-	-	-	
	[Hatt et al. 2015]	Cancers de l'œsophage	-		-	-	-	-			-		-
	[Hatt et al. 2015]	Cancers pulmonaires non à petites cellules	-		-	-	-	-			-		-
	[Tixier et al. 2014]	Cancers pulmonaires non à petites cellules			-	-	-				-		
	[Lapa et al. 2015]	Cancers de la thyroïde	-	*	*	-	-	*			-	-	-
Prédiction de la récurrence	[Pyka et al. 2015]	Cancers pulmonaires non à petites cellules	-	*	-	-	-	-			-		-

Tableau 5.6 – Tableau résumant les résultats présentés dans la littérature. Vert : p-value très inférieure à 5%. Jaune : p-value inférieure à 5%. Rouge : p-value supérieure à 5%. * signifie que p-value (index de texture) < p-value (index conventionnels).

	Homogeneity	Entropy	SRE	LRE	LGZE	HGZE	SUVmax	SUVmean	SUVpeak
Algorithme de reconstruction									
Nombre d'itérations									
Filtrage									
Taille de la grille de voxels									

Tableau 5.7 – Tableau résumant les résultats présentés dans [Yan et al., 2015]. Vert : plus ou aussi robuste que SUVmean ou SUVpeak. Jaune : moins robuste que SUVmean ou SUVpeak mais plus ou aussi robuste que SUVmax. Rouge : moins robuste que les mesures SUV.

Conclusion et perspectives

Actuellement, la prise en charge de patients atteints de cancer tend à être de plus en plus individualisée, en proposant un traitement spécifique en fonction des caractéristiques de la tumeur. Dans cette optique, l'identification de biomarqueurs s'avère indispensable pour établir le profil le plus précis possible de la lésion et suivre son évolution au cours du traitement. Les techniques d'imagerie fonctionnelle telles que l'imagerie TEP sont de bonnes candidates pour rechercher ces biomarqueurs. L'extraction de nouveaux paramètres à partir des images à l'avantage d'être non invasive, sur des examens déjà réalisés pendant le parcours de soin du patient, pour la lésion en entier, voire pour l'ensemble des lésions, et peut être répétée au cours du traitement.

Un nombre croissant de publications s'intéressent ainsi à l'analyse de texture à partir des images TEP afin d'identifier des paramètres prédictifs de la réponse à la thérapie ou de la survie du patient. Toutefois, l'état de l'art de la littérature a mis en lumière de grandes disparités entre les études : des protocoles de calcul différents ou non détaillés, ainsi que des conclusions parfois contradictoires entre les auteurs.

Avant de tester l'apport en clinique des index de texture, il nous est d'abord apparu indispensable d'étudier et de comprendre le rôle des différentes étapes de calcul de ces paramètres. Ainsi, nous avons montré que les index de texture pouvaient être réunis en groupes de corrélation, ne fournissant pas tous une information indépendante. L'identification de ces groupes a permis d'expliquer de nombreux résultats de la littérature et de diminuer le nombre de paramètres à calculer, limitant ainsi le nombre de résultats faux positifs.

L'étape de discrétisation des intensités dans le VOI est indispensable avant le calcul des index de texture afin de réduire la taille des matrices à calculer et de regrouper les voxels ayant des intensités proches. Dans environ 60% des publications, cette étape n'est pas documentée et dans un tiers des études, le taux de discrétisation n'est pas spécifié. Or, nous avons démontré que cette étape influençait fortement les valeurs des paramètres de texture. En effet, l'utilisation de la discrétisation relative dans la majorité des publications entraîne une corrélation entre les index de texture et VM du fait d'un artefact de calcul. La méthode alternative que nous proposons, la discrétisation absolue, permet de limiter cette corrélation et conserve l'information relative à l'intensité de fixation dans le calcul des index. Concernant le taux de discrétisation, nous avons montré qu'une valeur minimale de 32 niveaux est nécessaire et conseillons, pour la suite, l'utilisation de 64 niveaux, comme dans la plupart des publications rapportant cette étape.

Puisqu'aucune méthode de segmentation des lésions en imagerie TEP ne fait consensus, il s'est avéré nécessaire d'étudier son impact sur les valeurs des index de texture. Nous avons ainsi pu choisir, pour chaque groupe, un index parmi les plus robustes à la méthode de segmentation : Homogeneity, Entropy, SRE, LRE, LGZE et HGZE, à calculer en complément des index conventionnels. La réalisation de simulations nous a permis de distinguer deux groupes d'index : ceux qui sont plutôt sensibles aux macro-hétérogénéités et ceux qui sont plutôt influencés par les micro-hétérogénéités.

Une fois ces recommandations établies sur le calcul et le choix des index de texture, nous avons étudié leur capacité à distinguer différents types de tissus. L'analyse de texture permet de séparer le tissu tumoral du tissu hépatique et du tissu controlatéral. Les valeurs des index de texture sont également corrélées à l'analyse visuelle de l'hétérogénéité de fixation du traceur dans la tumeur évaluée par deux experts. Concernant les cancers mammaires, les index de texture sont sensibles à la présence de récepteurs aux œstrogènes et identifient les lésions triple-négatives (des tumeurs de mauvais pronostic). Dans les cancers pulmonaires, les index sont sensibles aux sous-types de cancers, au niveau de différenciation des lésions et à la présence d'une nécrose histologique. Enfin, dans les gliomes nouvellement diagnostiqués, la combinaison de Entropy et SUVmax mesurés sur des images TEP après injection de 18F-LDOPA permet de distinguer les lésions de haut grade des lésions de bas grade, et, dans les gliomes de haut grade, la texture des lésions nouvellement diagnostiquées est différente de celle des tumeurs en récurrence.

Finalement, une étude multi-échelle a été réalisée afin d'évaluer le potentiel et les limites de l'analyse de texture des images TEP. Nous avons montré que les index étaient modérément affectés par la taille des voxels ou la résolution spatiale des images. Sur des modèles animaux, la texture mesurée sur les images TEP *in vivo* reflète bien la texture mesurée *ex vivo* sur les images autoradiographiques. Même si les index calculés à partir des images TEP ne reflètent pas précisément la distribution de cellules observée sur les coupes histologiques, les paramètres sont sensibles, soit à la proportion de cellules tumorales dans le VOI, soit à l'hétérogénéité biologique du tissu.

À l'issue de ces analyses, parmi les 6 index sélectionnés, Entropy apparaît comme l'index conduisant aux résultats les plus prometteurs pour une utilisation en clinique. Il est très robuste vis-à-vis de la méthode de segmentation, modérément affecté par la taille des voxels et peu influencé par les différences de résolution spatiale. Entropy, qui correspond à une mesure du désordre, est capable de distinguer le tissu tumoral du tissu sain, est bien corrélé à l'analyse visuelle de l'hétérogénéité de fixation du traceur dans le volume tumoral, et permet notamment de distinguer les sous-types de cancers pulmonaires, le niveau de différenciation et la présence de nécrose histologique avec de meilleures performances que les index conventionnels (mesures SUV, VM ou TLG).

Ce travail nous a permis d'identifier de nombreuses perspectives pour approfondir les analyses réalisées.

Dans la dernière partie de ce manuscrit, nous avons présenté une analyse multi-échelle de l'hétérogénéité tumorale à partir d'images TEP au 18F-FDG et d'un marquage à l'hématoxyline des noyaux des cellules. Cette étude a montré que les index de texture étaient sensibles à l'hétérogénéité biologique sous-jacente. Néanmoins, ce travail préliminaire nécessite d'être complété avec l'utilisation à la fois d'autres traceurs en imagerie TEP et d'autres marquages des coupes histologiques afin de comprendre davantage la signification biologique de la texture mesurée *in vivo*. L'étude de l'angiogenèse, des zones hypoxiques ou de la présence de certains récepteurs permettrait de mettre en lumière quelles caractéristiques tumorales sont quantifiables à partir des examens TEP.

Pour valider l'utilisation de l'analyse de texture à partir des images TEP, la standardisation des méthodes de calcul des index s'avère indispensable, notamment en ce qui concerne la discrétisation absolue. Comme cela a été fait pour les mesures SUV [Zincirkeser et al., 2007], la construction de tables de référence correspondant à la texture calculée dans différents organes donnerait également aux cliniciens des valeurs de référence pour l'interprétation des paramètres. L'étude que nous avons amorcée ne porte pour l'instant que sur les valeurs des index dans le tissu hépatique, le tissu mammaire sain ou le tissu pulmonaire sain, mais la texture d'autres organes pourrait être calculée comme

pour la vessie, les reins ou les muscles. Les performances de l'analyse de texture pourraient également être évaluées pour différencier les processus inflammatoires des processus tumoraux, comme par exemple dans le cas des sarcoïdoses, ou pour quantifier spécifiquement les différents types d'hétérogénéité. L'hétérogénéité intra-tumorale peut aussi être quantifiée localement en réalisant des cartes de texture où les index sont calculés dans des « fenêtres » (par exemple de $5 \times 5 \times 5$ voxels) et où la valeur du paramètre de texture est affectée au voxel central. Cette cartographie pourrait permettre de guider la biopsie en identifiant la ou les zones les plus hétérogènes ainsi qu'aider à la planification de traitement en radiothérapie. L'hétérogénéité inter-tumorale pourrait aussi être analysée en comparant, pour un même patient, la texture entre la tumeur primitive et les différentes métastases.

Afin de confirmer l'intérêt du calcul des index de texture à partir des images TEP, des méta-analyses vont être nécessaires pour recueillir suffisamment de preuves de leur valeur informative, avant leur intégration dans le parcours de soin du patient. La mesure des index de texture ne remet pas en cause l'utilisation des index conventionnels mais doit être vue comme un complément aux index déjà existants. Une démonstration claire de l'intérêt de ces paramètres et l'identification d'applications spécifiques sont indispensables pour pérenniser l'utilisation de l'analyse de texture en imagerie TEP. En effet, l'évaluation de l'hétérogénéité tumorale a commencé par le calcul d'index de texture à partir d'images échographiques, notamment dans le cas des cancers mammaires [Garra et al., 1993]. Or, ces paramètres de texture ne sont actuellement pas utilisés en routine clinique dans cette modalité, une vingtaine d'années après la publication des premiers articles. Comme le souligne [Alic et al., 2014], très peu d'études prospectives ont été menées à partir des images échographiques. De plus, la majorité des publications portait sur la capacité de l'analyse de texture à évaluer le type ou le grade de la tumeur, alors qu'en pratique clinique le recours à la biopsie pré-traitement fournit déjà ces informations. À court terme, les index de texture issus des images ne doivent pas être vus comme un substitut de la biopsie initiale, quand elle est réalisable, mais comme un complément en fournissant des informations additionnelles et possiblement à plusieurs reprises au cours du traitement.

Pour quantifier encore plus précisément l'hétérogénéité tumorale visualisée par les médecins nucléaires lors de l'évaluation des images TEP, de nouveaux index pourraient aussi être proposés, par exemple pour rendre compte de la position de la nécrose ou des clones hypermétaboliques. En effet, au moyen des simulations, nous avons montré que les index de texture calculés actuellement n'étaient pas sensibles à la position des hétérogénéités dans le volume tumoral. Ces nouveaux index, en complément de ceux déjà existants, pourraient permettre d'identifier plus systématiquement et plus précocement l'organisation spatiale de la tumeur.

Dans le concept initial de radiomique, les paramètres peuvent être extraits de différentes modalités d'imagerie. Cependant, avant de généraliser l'analyse de texture à la multi-modalité, il est nécessaire de mener une étude similaire à celle que nous avons réalisée sur les images TEP afin de déterminer la procédure la plus adéquate pour les calculs des paramètres, en particulier en ce qui concerne l'étape de discrétisation, et de comprendre mieux la signification biologique de ces index. Une co-analyse des textures observées dans différentes modalités pourrait permettre de combiner des informations complémentaires. Par exemple, pour les gliomes, les données fournies par un examen IRM et un examen TEP au 18F-FLT, avant traitement, permettent déjà une bonne classification des lésions à partir du calcul d'index conventionnels [Collet et al., 2015]. L'ajout de paramètres plus complexes permettrait peut-être d'améliorer encore davantage les performances de classification. Une première étude dans ce sens a été réalisée par [Vallières et al., 2015] qui combinent à la fois des index de texture, de forme ou de l'histogramme issus des images TEP au 18F-FDG et des images IRM afin de prédire l'apparition de métastases pulmonaires dans des sarcomes des tissus mous. Ce type d'analyse bi-modalité va notamment se développer avec l'installation de machines hybrides TEP-IRM permettant d'obtenir des images recalées et simultanées à partir d'un seul examen. L'utilisation de plusieurs traceurs

pour un même patient devrait également permettre de caractériser plus précisément l'hétérogénéité biologique des tumeurs en quantifiant, par exemple, la prolifération cellulaire ou l'apoptose en plus du métabolisme glucidique.

À l'heure actuelle, aucun logiciel ne permet le calcul de l'ensemble des paramètres de texture quelque soit le type d'images. Par exemple, le logiciel TexRad [[Ganeshan et al., 2010](#)], contrairement à ce que pourrait évoquer son nom, n'offre pas la possibilité de calculer les index de texture, mais uniquement les paramètres issus de l'histogramme des niveaux de gris. Le logiciel MaZda [[Szczypiński et al., 2009](#)], développé à l'origine pour les images IRM, ne permet pas de calculer les index issus de GLZLM, les paramètres propres aux images TEP comme SUVpeak ou TLG, ainsi que de discrétiser les intensités avec la méthode absolue. La disponibilité d'un outil complet, compatible avec plusieurs modalités d'imagerie, incluant les index issus de l'histogramme, les index de forme, les index de texture, les méthodes de décomposition en ondelette ou d'analyse fractale, permettrait de développer davantage l'étude et l'utilisation de biomarqueurs issus des examens d'imagerie.

À terme, le concept de radiomique présente de nombreuses applications notamment pour la planification de traitement en radiothérapie. En effet, environ 50% des patients atteints d'un cancer bénéficieront d'un traitement par radiothérapie. Or, cette technique tend également à mieux prendre en compte les spécificités de la tumeur pour guider la stratégie thérapeutique [[Bibault et al., 2013](#)]. À partir des examens d'imagerie, les paramètres extraits pourraient permettre d'identifier, avant le début du traitement, les lésions potentiellement radio-résistantes. Aussi, une mesure de l'hétérogénéité locale, à l'échelle du voxel, à partir de cartes de texture par exemple, pourrait offrir la possibilité d'adapter le plan de traitement et la répartition du rayonnement délivré.

Une autre application est portée par le projet QuI-ConCEPT [[Waterton and Pylkkanen, 2012](#)]. Ce consortium vise à développer trois biomarqueurs issus des données d'imagerie pour quantifier la prolifération cellulaire, l'apoptose et la nécrose afin de démontrer l'efficacité d'un médicament durant les essais cliniques de Phase I, de manière simple et robuste pour des études multicentriques. À terme, ces développements et les connaissances acquises durant ce programme bénéficieront à la prise en charge des patients en routine clinique.

Bibliographie

- [Aerts et al., 2009] Aerts, H. J. W. L., van Baardwijk, A. A. W., Petit, S. F., Offermann, C., Loon, J., Houben, R., Dingemans, A.-M. C., Wanders, R., Boersma, L., Borger, J., Bootsma, G., Geraedts, W., Pitz, C., Simons, J., Wouters, B. G., Oellers, M., Lambin, P., Bosmans, G., Dekker, A. L. A. J., and De Ruyscher, D. (2009). Identification of residual metabolic-active areas within individual NSCLC tumours using a pre-radiotherapy (18)Fluorodeoxyglucose-PET-CT scan. *Radiother Oncol*, 91(3):386–392. [28](#)
- [Ahmed et al., 2013] Ahmed, A., Gibbs, P., Pickles, M., and Turnbull, L. (2013). Texture analysis in assessment and prediction of chemotherapy response in breast cancer. *J Magn Reson Imaging*, 38(1):89–101. [119](#)
- [Alic et al., 2014] Alic, L., Niessen, W. J., and Veenland, J. F. (2014). Quantification of heterogeneity as a biomarker in tumor imaging: a systematic review. *PloS One*, 9(10):e110300. [26](#), [165](#)
- [Amadasun and King, 1989] Amadasun, M. and King, R. (1989). Textural features corresponding to textural properties. *IEEE Trans Syst Man Cybern*, 19(5):1264–1274. [40](#)
- [Apostolova et al., 2014a] Apostolova, I., Rogasch, J., Buchert, R., Wertz, H., Achenbach, H. J., Schreiber, J., Riedel, S., Furth, C., Lougovski, A., Schramm, G., Hofheinz, F., Amthauer, H., and Steffen, I. G. (2014a). Quantitative assessment of the asphericity of pretherapeutic FDG uptake as an independent predictor of outcome in NSCLC. *BMC cancer*, 14:896. [26](#)
- [Apostolova et al., 2014b] Apostolova, I., Steffen, I. G., Wedel, F., Lougovski, A., Marnitz, S., Derlin, T., Amthauer, H., Buchert, R., Hofheinz, F., and Brenner, W. (2014b). Asphericity of pretherapeutic tumour FDG uptake provides independent prognostic value in head-and-neck cancer. *Eur Radiol*, 24(9):2077–2087. [26](#)
- [Aristophanous et al., 2007] Aristophanous, M., Penney, B. C., Martel, M. K., and Pelizzari, C. A. (2007). A Gaussian mixture model for definition of lung tumor volumes in positron emission tomography. *Med Phys*, 34(11):4223–4235. [21](#)
- [Asamura et al., 2008] Asamura, H., Goya, T., Koshiishi, Y., Sohara, Y., Eguchi, K., Mori, M., Nakanishi, Y., Tsuchiya, R., Shimokata, K., Inoue, H., Nukiwa, T., Miyaoka, E., and Japanese Joint Committee of Lung Cancer Registry (2008). A Japanese lung cancer registry study: prognosis of 13,010 resected lung cancers. *J Thorac Oncol*, 3(1):46–52. [125](#)
- [Asselin et al., 2012] Asselin, M.-C., O’Connor, J. P. B., Boellaard, R., Thacker, N. A., and Jackson, A. (2012). Quantifying heterogeneity in human tumours using MRI and PET. *Eur J Cancer*, 48(4):447–455. [44](#)
- [Bagci et al., 2013] Bagci, U., Yao, J., Miller-Jaster, K., Chen, X., and Mollura, D. J. (2013). Predicting future morphological changes of lesions from radiotracer uptake in 18F-FDG-PET images. *PloS One*, 8(2):e57105. [46](#)
- [Bailey et al., 2005] Bailey, D. L., Townsend, D. W., Valk, P. E., and Maisey, M. N. (2005). *Positron Emission Tomography*. Springer-Verlag, London. [8](#)
- [Beucher and Meyer, 1992] Beucher, S. and Meyer, F. (1992). The morphological approach to segmentation: the watershed transformation. *Opt Eng*, 34(12):433. [20](#)

- [Beyer et al., 2011] Beyer, T., Czernin, J., and Freudenberg, L. S. (2011). Variations in clinical PET/CT operations: results of an international survey of active PET/CT users. *J Nucl Med*, 52(2):303–310. [22](#)
- [Bibault et al., 2013] Bibault, J.-E., Fumagalli, I., Fertié, C., Chargari, C., Soria, J.-C., and Deutsch, E. (2013). Personalized radiation therapy and biomarker-driven treatment strategies: a systematic review. *Cancer Metastasis Rev*, 32(3):479–492. [166](#)
- [Billaud, 2012] Billaud, M. (2012). L’hétérogénéité intratumorale. *Med Sci*, 28:1116–1119. [34](#)
- [Bollineni et al., 2014] Bollineni, V. R., Koole, M. J. B., Pruim, J., Brouwer, C. L., Wiegman, E. M., Groen, H. J. M., Vlasman, R., Halmos, G. B., Oosting, S. F., Langendijk, J. A., Widder, J., and Steenbakkers, R. J. H. M. (2014). Dynamics of tumor hypoxia assessed by 18F-FAZA PET/CT in head and neck and lung cancer patients during chemoradiation: possible implications for radiotherapy treatment planning strategies. *Radioth Oncol*, 113(2):198–203. [28](#)
- [Brooks, 2013] Brooks, F. J. (2013). Area under the cumulative SUV-volume histogram is not a viable metric of intratumoral metabolic heterogeneity: further comments. *Eur J Nucl Med Mol Imaging*, 40(12):1926–1927. [25](#)
- [Brooks and Grigsby, 2014] Brooks, F. J. and Grigsby, P. W. (2014). The effect of small tumor volumes on studies of intratumoral heterogeneity of tracer uptake. *J Nucl Med*, 55(1):37–42. [46](#), [49](#), [75](#), [77](#), [78](#), [157](#), [159](#)
- [Brooks and Grigsby, 2015] Brooks, F. J. and Grigsby, P. W. (2015). Low-order non-spatial effects dominate second-order spatial effects in the texture quantifier analysis of 18F-FDG-PET images. *PloS One*, 10(2):e0116574. [47](#), [49](#), [135](#), [157](#)
- [Browne and de Pierro, 1996] Browne, J. and de Pierro, A. B. (1996). A row-action alternative to the EM algorithm for maximizing likelihood in emission tomography. *IEEE Trans Med Imaging*, 15(5):687–699. [11](#)
- [Bundschuh et al., 2014] Bundschuh, R. A., Dinges, J., Neumann, L., Seyfried, M., Zsótér, N., Papp, L., Rosenberg, R., Becker, K., Astner, S. T., Henninger, M., Herrmann, K., Ziegler, S. I., Schwaiger, M., and Essler, M. (2014). Textural parameters of tumor heterogeneity in 18F-FDG PET/CT for therapy response assessment and prognosis in patients with locally advanced rectal cancer. *J Nucl Med*, 55(6):891–897. [23](#)
- [Buvat, 2002] Buvat, I. (2002). A non-parametric bootstrap approach for analysing the statistical properties of SPECT and PET images. *Phys Med Biol*, 47(10):1761–1775. [77](#)
- [Buvat, 2007a] Buvat, I. (2007a). Les limites du SUV. *Med Nucl*, 31(4):165–172. [17](#), [19](#)
- [Buvat, 2007b] Buvat, I. (2007b). Quantification in emission tomography: Challenges, solutions, and performance. *Nucl Instr Meth Phys Res*, 571(1–2):10–13. [12](#)
- [Buvat et al., 2012] Buvat, I., Necib, H., Garcia, C., Wagner, A., Vanderlinden, B., Emonts, P., Hendlisz, A., and Flamen, P. (2012). Lesion-based detection of early chemosensitivity using serial static FDG PET/CT in metastatic colorectal cancer. *Eur J Nucl Med Mol Imaging*, 39(10):1628–1634. [57](#)
- [Carson, 2005] Carson, R. E. (2005). Tracer Kinetic Modeling in PET. In *Positron Emission Tomography*, pages 127–159. Springer London. [15](#)
- [Chalkidou et al., 2015] Chalkidou, A., O’Doherty, M. J., and Marsden, P. K. (2015). False discovery rates in PET and CT studies with texture features: a systematic review. *PloS One*, 10(5):e0124165. [56](#), [92](#), [95](#), [157](#)
- [Chen et al., 2001] Chen, J. L., Gunn, S. R., Nixon, M. S., and Gunn, R. N. (2001). Markov random field models for segmentation of PET images. *Information Processing in Medical Imaging, Lect Notes Comput*, 2082:468–474. [21](#)

- [Cheng et al., 2013] Cheng, N.-M., Fang, Y.-H. D., Chang, J. T.-C., Huang, C.-G., Tsan, D.-L., Ng, S.-H., Wang, H.-M., Lin, C.-Y., Liao, C.-T., and Yen, T.-C. (2013). Textural features of pretreatment 18F-FDG PET/CT images: prognostic significance in patients with advanced T-stage oropharyngeal squamous cell carcinoma. *J Nucl Med*, 54(10):1703–1709. [45](#), [54](#), [159](#)
- [Cheng et al., 2015] Cheng, N.-M., Fang, Y.-H. D., Lee, L.-Y., Chang, J. T.-C., Tsan, D.-L., Ng, S.-H., Wang, H.-M., Liao, C.-T., Yang, L.-Y., Hsu, C.-H., and Yen, T.-C. (2015). Zone-size nonuniformity of 18F-FDG PET regional textural features predicts survival in patients with oropharyngeal cancer. *Eur J Nucl Med Mol Imaging*, 42(3):419–428. [47](#), [52](#), [54](#)
- [Chicklore et al., 2013] Chicklore, S., Goh, V., Siddique, M., Roy, A., Marsden, P. K., and Cook, G. J. R. (2013). Quantifying tumour heterogeneity in 18F-FDG PET/CT imaging by texture analysis. *Eur J Nucl Med Mol Imaging*, 40(1):133–140. [44](#)
- [Collet et al., 2015] Collet, S., Valable, S., Constans, J., Lechapt-Zalcman, E., Roussel, S., Delcroix, N., Abbas, A., Ibazizene, M., Bernaudin, M., Barré, L., Derlon, J., and Guillamo, J. (2015). [18F]-fluoro-l-thymidine PET and advanced MRI for preoperative grading of gliomas. *Neuroimage Clin*, 8:448–454. [165](#)
- [Cook et al., 2015] Cook, G. J. R., O'Brien, M. E., Siddique, M., Chicklore, S., Loi, H. Y., Sharma, B., Punwani, R., Bassett, P., Goh, V., and Chua, S. (2015). Non-small cell lung cancer treated with Erlotinib: heterogeneity of (18)F-FDG uptake at PET-association with treatment response and prognosis. *Radiology*, 276(3):883–893. [47](#), [54](#), [67](#)
- [Cook et al., 2013] Cook, G. J. R., Yip, C., Siddique, M., Goh, V., Chicklore, S., Roy, A., Marsden, P., Ahmad, S., and Landau, D. (2013). Are pretreatment 18F-FDG PET tumor textural features in non-small cell lung cancer associated with response and survival after chemoradiotherapy? *J Nucl Med*, 54(1):19–26. [46](#), [54](#), [55](#), [67](#)
- [Darcourt et al., 2014] Darcourt, J., Schiazza, A., Sapin, N., Dufour, M., Ouvrier, M. J., Benisvy, D., Fontana, X., and Koulibaly, P. M. (2014). 18F-FDOPA PET for the diagnosis of parkinsonian syndromes. *Q J Nucl Med Mol Imaging*, 58(4):355–365. [14](#)
- [Davnall et al., 2012] Davnall, F., Yip, C. S. P., Ljungqvist, G., Selmi, M., Ng, F., Sanghera, B., Ganeshan, B., Miles, K. A., Cook, G. J., and Goh, V. (2012). Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice? *Insights Imaging*, 3(6):573–589. [44](#)
- [Doll et al., 2014] Doll, C., Duncker-Rohr, V., Rücker, G., Mix, M., MacManus, M., De Ruyscher, D., Vogel, W., Eriksen, J. G., Oyen, W., Grosu, A.-L., Weber, W., and Nestle, U. (2014). Influence of experience and qualification on PET-based target volume delineation. When there is no expert—ask your colleague. *Strahlenther Onkol*, 190(6):555–562. [28](#)
- [Dong et al., 2015] Dong, X., Wu, P., Sun, X., Li, W., Wan, H., Yu, J., and Xing, L. (2015). Intratumour 18F-FDG uptake heterogeneity decreases the reliability on target volume definition with positron emission tomography/computed tomography imaging. *J Med Imaging Rad Oncol*, 59(3):338–345. [47](#), [55](#)
- [Dong et al., 2013] Dong, X., Xing, L., Wu, P., Fu, Z., Wan, H., Li, D., Yin, Y., Sun, X., and Yu, J. (2013). Three-dimensional positron emission tomography image texture analysis of esophageal squamous cell carcinoma: relationship between tumor 18F-fluorodeoxyglucose uptake heterogeneity, maximum standardized uptake value, and tumor stage. *Nucl Med Comm*, 34(1):40–46. [46](#), [54](#), [61](#), [63](#), [92](#), [125](#), [159](#)
- [Drever et al., 2007] Drever, L., Roa, W., McEwan, A., and Robinson, D. (2007). Iterative threshold segmentation for PET target volume delineation. *Med Phys*, 34(4):1253–1265. [20](#)
- [Eary et al., 2008] Eary, J. F., O'Sullivan, F., O'Sullivan, J., and Conrad, E. U. (2008). Spatial heterogeneity in sarcoma 18F-FDG uptake as a predictor of patient outcome. *J Nucl Med*, 49(12):1973–1979. [26](#)

- [Eisenhauer et al., 2009] Eisenhauer, E. A., Therasse, P., Bogaerts, J., Schwartz, L. H., Sargent, D., Ford, R., Dancey, J., Arbuck, S., Gwyther, S., Mooney, M., Rubinstein, L., Shankar, L., Dodd, L., Kaplan, R., Lacombe, D., and Verweij, J. (2009). New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*, 45(2):228–247. [29](#)
- [El Naqa et al., 2009] El Naqa, I., Grigsby, P., Apte, A., Kidd, E., Donnelly, E., Khullar, D., Chaudhari, S., Yang, D., Schmitt, M., Laforest, R., Thorstad, W., and Deasy, J. O. (2009). Exploring feature-based approaches in PET images for predicting cancer treatment outcomes. *Pattern Recogn*, 42(6):1162–1171. [25](#), [26](#), [45](#), [48](#), [54](#), [63](#), [92](#), [159](#), [160](#)
- [Erlandsson et al., 2012] Erlandsson, K., Buvat, I., Pretorius, P. H., Thomas, B. A., and Hutton, B. F. (2012). A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology. *Phys Med Biol*, 57(21):R119–159. [19](#)
- [Fornier and Fumoleau, 2012] Fornier, M. and Fumoleau, P. (2012). The paradox of triple negative breast cancer: novel approaches to treatment. *Breast J*, 18(1):41–51. [119](#)
- [Foster et al., 2014] Foster, B., Bagci, U., Mansoor, A., Xu, Z., and Mollura, D. J. (2014). A review on segmentation of positron emission tomography images. *Comput Biol Med*, 50:76–96. [20](#)
- [Galavis et al., 2010] Galavis, P. E., Hollensen, C., Jallow, N., Paliwal, B., and Jeraj, R. (2010). Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters. *Acta Oncol*, 49(7):1012–1016. [45](#), [50](#), [51](#), [67](#), [159](#), [160](#)
- [Ganeshan et al., 2010] Ganeshan, B., Abaleke, S., Young, R. C. D., Chatwin, C. R., and Miles, K. A. (2010). Texture analysis of non-small cell lung cancer on unenhanced computed tomography: initial evidence for a relationship with tumour glucose metabolism and stage. *Cancer Imaging*, 10:137–143. [166](#)
- [Gao et al., 2015] Gao, X., Chu, C., Li, Y., Lu, P., Wang, W., Liu, W., and Yu, L. (2015). The method and efficacy of support vector machine classifiers based on texture features and multi-resolution histogram from (18)F-FDG PET-CT images for the evaluation of mediastinal lymph nodes in patients with lung cancer. *Eur J Radiol*, 84(2):312–317. [47](#)
- [Garra et al., 1993] Garra, B. S., Krasner, B. H., Horii, S. C., Ascher, S., Mun, S. K., and Zeman, R. K. (1993). Improving the distinction between benign and malignant breast lesions: the value of sonographic texture analysis. *Ultrason Imaging*, 15(4):267–285. [165](#)
- [Gebejes and Huertas, 2013] Gebejes, A. and Huertas, R. (2013). Texture characterization based on grey-level co-occurrence matrix. *Conference of Informatics and Management Sciences, March 25–29*, 2(1):375–378. [91](#)
- [Gerlinger et al., 2012] Gerlinger, M., Rowan, A. J., Horswell, S., Larkin, J., Endesfelder, D., Gronroos, E., Martinez, P., Matthews, N., Stewart, A., Tarpey, P., Varela, I., Phillimore, B., Begum, S., McDonald, N. Q., Butler, A., Jones, D., Raine, K., Latimer, C., Santos, C. R., Nohadani, M., Eklund, A. C., Spencer-Dene, B., Clark, G., Pickering, L., Stamp, G., Gore, M., Szallasi, Z., Downward, J., Futreal, P. A., and Swanton, C. (2012). Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med*, 366(10):883–892. [35](#)
- [Hall, 1971] Hall, R. (1971). Vascular injuries resulting from arterial puncture of catheterization. *Br J Surg*, 58(7):513–516. [16](#)
- [Hamberg et al., 1994] Hamberg, L. M., Hunter, G. J., Alpert, N. M., Choi, N. C., Babich, J. W., and Fischman, A. J. (1994). The dose uptake ratio as an index of glucose metabolism: useful parameter or oversimplification? *J Nucl Med*, 35(8):1308–1312. [18](#)
- [Hanahan and Weinberg, 2000] Hanahan, D. and Weinberg, R. A. (2000). The hallmarks of cancer. *Cell*, 100(1):57–70. [33](#)
- [Hanahan and Weinberg, 2011] Hanahan, D. and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell*, 144(5):646–674. [33](#), [34](#)

- [Hapdey et al., 2011] Hapdey, S., Buvat, I., Carson, J. M., Carson, J. M., Carrasquillo, J. A., Whatley, M., and Bacharach, S. L. (2011). Searching for alternatives to full kinetic analysis in 18F-FDG PET: an extension of the simplified kinetic analysis method. *J Nucl Med*, 52(4):634–641. [17](#)
- [Haralick et al., 1973] Haralick, R., Shanmugam, K., and Dinstein, I. (1973). Textural features for image classification. *IEEE Trans Syst Man Cybern*, 3(6):610–621. [37](#)
- [Hatt et al., 2009] Hatt, M., Cheze le Rest, C., Turzo, A., Roux, C., and Visvikis, D. (2009). A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET. *IEEE Trans Med Imaging*, 28(6):881–893. [21](#)
- [Hatt et al., 2015] Hatt, M., Majdoub, M., Vallières, M., Tixier, F., Cheze Le Rest, C., Groheux, D., Hindié, E., Martineau, A., Pradier, O., Hustinx, R., Perdrisot, R., Guillevin, R., El Naqa, I., and Visvikis, D. (2015). 18F-FDG PET uptake characterization through texture analysis: investigating the complementary nature of heterogeneity and functional tumor volume in a multi-cancer site patient cohort. *J Nucl Med*, 56(1):38–44. [47](#), [48](#), [49](#), [54](#), [92](#), [126](#), [159](#)
- [Hatt et al., 2013] Hatt, M., Tixier, F., Cheze Le Rest, C., Pradier, O., and Visvikis, D. (2013). Robustness of intratumour 18F-FDG PET uptake heterogeneity quantification for therapy response prediction in oesophageal carcinoma. *Eur J Nucl Med Mol Imaging*, 40(11):1662–1671. [45](#), [50](#), [51](#), [52](#), [53](#), [61](#), [62](#), [64](#), [92](#), [159](#)
- [Hatt et al., 2011] Hatt, M., Visvikis, D., Pradier, O., and Cheze le Rest, C. (2011). Baseline 18F-FDG PET image-derived parameters for therapy response prediction in oesophageal cancer. *Eur J Nucl Med Mol Imaging*, 38(9):1595–1606. [61](#), [157](#)
- [Hendlisz et al., 2012] Hendlisz, A., Golfopoulos, V., Garcia, C., Covas, A., Emonts, P., Ameye, L., Paesmans, M., Deleporte, A., Machiels, G., Toussaint, E., Vanderlinden, B., Awada, A., Piccart, M., and Flamen, P. (2012). Serial FDG-PET/CT for early outcome prediction in patients with metastatic colorectal cancer undergoing chemotherapy. *Ann Oncol*, 23(7):1687–1693. [57](#)
- [Hoeben et al., 2014] Hoeben, B. A. W., Starmans, M. H. W., Leijenaar, R. T. H., Dubois, L. J., van der Kogel, A. J., Kaanders, J. H. A. M., Boutros, P. C., Lambin, P., and Bussink, J. (2014). Systematic analysis of 18F-FDG PET and metabolism, proliferation and hypoxia markers for classification of head and neck tumors. *BMC cancer*, 14:130. [150](#)
- [Hoekstra et al., 2002] Hoekstra, C. J., Hoekstra, O. S., Stroobants, S. G., Vansteenkiste, J., Nuyts, J., Smit, E. F., Boers, M., Twisk, J. W. R., and Lammertsma, A. A. (2002). Methods to monitor response to chemotherapy in non-small cell lung cancer with 18F-FDG PET. *J Nucl Med*, 43(10):1304–1309. [17](#)
- [Hofheinz et al., 2015] Hofheinz, F., Lougovski, A., Zöphel, K., Hentschel, M., Steffen, I. G., Apostolova, I., Wedel, F., Buchert, R., Baumann, M., Brenner, W., Kotzerke, J., and van den Hoff, J. (2015). Increased evidence for the prognostic value of primary tumor asphericity in pretherapeutic FDG PET for risk stratification in patients with head and neck cancer. *Eur J Nucl Med Mol Imaging*, 42(3):429–437. [26](#)
- [Hofmann et al., 2009] Hofmann, M., Pichler, B., Schölkopf, B., and Beyer, T. (2009). Towards quantitative PET/MRI: a review of MR-based attenuation correction techniques. *Eur J Nucl Med Mol Imaging*, 36 Suppl 1:S93–104. [12](#)
- [Holm, 1979] Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand J Stat*, 6:65–70. [97](#)
- [Houshmand et al., 2015] Houshmand, S., Salavati, A., Hess, S., Werner, T. J., Alavi, A., and Zaidi, H. (2015). An update on novel quantitative techniques in the context of evolving whole-body PET imaging. *PET Clin*, 10(1):45–58. [44](#)
- [Hsu et al., 2008] Hsu, C.-Y., Liu, C.-Y., and Chen, C.-M. (2008). Automatic segmentation of liver PET images. *Comput Med Imaging Graph*, 32(7):601–610. [20](#)

- [Hudson and Larkin, 1994] Hudson, H. M. and Larkin, R. S. (1994). Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans Med Imaging*, 13(4):601–609. [11](#)
- [Hunter et al., 1996] Hunter, G. J., Hamberg, L. M., Alpert, N. M., Choi, N. C., and Fischman, A. J. (1996). Simplified measurement of deoxyglucose utilization rate. *J Nucl Med*, 37(6):950–955. [17](#)
- [Hutton, 2011] Hutton, B. F. (2011). Recent advances in iterative reconstruction for clinical SPECT/PET and CT. *Acta Oncol*, 50(6):851–858. [19](#)
- [Jentzen et al., 2007] Jentzen, W., Freudenberg, L., Eising, E. G., Heinze, M., Brandau, W., and Bockisch, A. (2007). Segmentation of PET volumes by iterative image thresholding. *J Nucl Med*, 48(1):108–114. [20](#)
- [Kang et al., 2014] Kang, S.-R., Song, H.-C., Byun, B. H., Oh, J.-R., Kim, H.-S., Hong, S.-P., Kwon, S. Y., Chong, A., Kim, J., Cho, S.-G., Park, H. J., Kim, Y.-C., Ahn, S.-J., Min, J.-J., and Bom, H.-S. (2014). Intratumoral metabolic heterogeneity for prediction of disease progression after concurrent chemoradiotherapy in patients with inoperable stage III non-small-cell lung cancer. *Nucl Med Mol Imaging*, 48(1):16–25. [25](#)
- [Kendi et al., 2015a] Kendi, A. T., Corey, A., Magliocca, K. R., Nickleach, D. C., Galt, J., Switchenko, J. M., El-Deiry, M. W., Wadsworth, J. T., Hudgins, P. A., Saba, N. F., and Schuster, D. M. (2015a). 18F-FDG-PET/CT parameters as imaging biomarkers in oral cavity squamous cell carcinoma, is visual analysis of PET and contrast enhanced CT better than the numbers? *Eur J Radiol*, 84(6):1171–1176. [24](#)
- [Kendi et al., 2015b] Kendi, A. T. K., Magliocca, K., Corey, A., Nickleach, D. C., Galt, J., Higgins, K., Beitler, J. J., El-Deiry, M. W., Wadsworth, J. T., Hudgins, P. A., Saba, N. F., and Schuster, D. M. (2015b). Do 18F-FDG PET/CT parameters in oropharyngeal and oral cavity squamous cell carcinomas indicate HPV status? *Clin Nucl Med*, 40(3):196–200. [24](#)
- [Kim et al., 1994] Kim, C. K., Gupta, N. C., Chandramouli, B., and Alavi, A. (1994). Standardized uptake values of FDG: body surface area correction is preferable to body weight correction. *J Nucl Med*, 35(1):164–167. [18](#)
- [Kinahan et al., 2003] Kinahan, P. E., Hasegawa, B. H., and Beyer, T. (2003). X-ray-based attenuation correction for positron emission tomography/computed tomography scanners. *Sem Nucl Med*, 33(3):166–179. [19](#)
- [Kumar et al., 2012] Kumar, V., Gu, Y., Basu, S., Berglund, A., Eschrich, S. A., Schabath, M. B., Forster, K., Aerts, H. J. W. L., Dekker, A., Fenstermacher, D., Goldgof, D. B., Hall, L. O., Lambin, P., Balagurunathan, Y., Gatenby, R. A., and Gillies, R. J. (2012). Radiomics: the process and the challenges. *Magn Res Imaging*, 30(9):1234–1248. [5](#), [44](#)
- [Lambin et al., 2012] Lambin, P., Rios-Velazquez, E., Leijenaar, R., Carvalho, S., van Stiphout, R. G. P. M., Granton, P., Zegers, C. M. L., Gillies, R., Boellard, R., Dekker, A., and Aerts, H. J. W. L. (2012). Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*, 48(4):441–446. [5](#)
- [Lapa et al., 2015] Lapa, C., Werner, R. A., Schmid, J.-S., Papp, L., Zsótér, N., Biko, J., Reiners, C., Herrmann, K., Buck, A. K., and Bundschuh, R. A. (2015). Prognostic value of positron emission tomography-assessed tumor heterogeneity in patients with thyroid cancer undergoing treatment with radiopeptide therapy. *Nucl Med Biol*, 42(4):349–354. [47](#), [54](#), [158](#), [159](#)
- [Larson et al., 1999] Larson, S. M., Erdi, Y., Akhurst, T., Mazumdar, M., Macapinlac, H. A., Finn, R. D., Casilla, C., Fazzari, M., Srivastava, N., Yeung, H. W. D., Humm, J. L., Guillem, J., Downey, R., Karpeh, M., Cohen, A. E., and Ginsberg, R. (1999). Tumor treatment response based on visual and quantitative changes in global tumor glycolysis using PET-FDG Imaging. The visual response score and the change in total lesion glycolysis. *Clin Positron Imaging*, 2(3):159–171. [22](#)

- [Lartizien et al., 2014] Lartizien, C., Rogez, M., Niaf, E., and Ricard, F. (2014). Computer-aided staging of lymphoma patients with FDG PET/CT imaging based on textural information. *IEEE J Biomed Health Inform*, 18(3):946–955. [46](#), [48](#), [54](#)
- [Leijenaar et al., 2013] Leijenaar, R. T. H., Carvalho, S., Velazquez, E. R., van Elmpt, W. J. C., Parmar, C., Hoekstra, O. S., Hoekstra, C. J., Boellaard, R., Dekker, A. L. A. J., Gillies, R. J., Aerts, H. J. W. L., and Lambin, P. (2013). Stability of FDG-PET Radiomics features: an integrated analysis of test-retest and inter-observer variability. *Acta Oncol*, 52(7):1391–1397. [45](#), [49](#), [50](#), [51](#), [52](#), [53](#), [66](#), [78](#), [91](#), [158](#), [159](#), [160](#)
- [MacManus et al., 2009] MacManus, M., Nestle, U., Rosenzweig, K. E., Carrio, I., Messa, C., Belohlavek, O., Danna, M., Inoue, T., Deniaud-Alexandre, E., Schipani, S., Watanabe, N., Dondi, M., and Jeremic, B. (2009). Use of PET and PET/CT for radiation therapy planning: IAEA expert report 2006-2007. *Radioth Oncol*, 91(1):85–94. [27](#)
- [MacQueen, 1967] MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. Western Management Science Inst, Univ of California. [21](#)
- [Maisonobe et al., 2013] Maisonobe, J.-A., Garcia, C. A., Necib, H., Vanderlinden, B., Hendlisz, A., Flamen, P., and Buvat, I. (2013). Comparison of PET metabolic indices for the early assessment of tumour response in metastatic colorectal cancer patients treated by polychemotherapy. *Eur J Nucl Med Mol Imaging*, 40(2):166–174. [59](#)
- [Mangili et al., 1998] Mangili, F., Cigala, C., Arrigoni, G., Rovere, E., Gattuso, C., Santambrogio, G., and Garancini, P. (1998). Cell loss and proliferation in non-small cell lung carcinoma: correlation with histological subtype. *Eur J Histochem*, 42(4):287–295. [125](#)
- [Mankoff et al., 2006] Mankoff, D. D. A., Muzi, D. M., and Zaidi, P. D. H. (2006). Quantitative analysis in nuclear oncologic imaging. In *Quantitative Analysis in Nuclear Medicine Imaging*, pages 494–536. Springer US. [16](#)
- [Markel et al., 2013] Markel, D., Caldwell, C., Alasti, H., Soliman, H., Ung, Y., Lee, J., and Sun, A. (2013). Automatic segmentation of lung carcinoma using 3D texture features in 18F-FDG PET/CT. *Int J Mol Imaging*, 2013:980769. [45](#), [55](#)
- [Marusyk and Polyak, 2010] Marusyk, A. and Polyak, K. (2010). Tumor heterogeneity: causes and consequences. *Biochim Biophys Acta*, 1805(1):105–117. [35](#)
- [McAuliffe et al., 2001] McAuliffe, M., Lalonde, F., McGarry, D., Gandler, W., Csaky, K., and Trus, B. (2001). Medical image processing, analysis and visualization in clinical research. In *14th IEEE Symp Comp Med Syst*, pages 381–386. [138](#)
- [Meijering et al., 2001] Meijering, E. H. W., Niessen, W. J., and Viergever, M. A. (2001). Quantitative evaluation of convolution-based methods for medical image interpolation. *Med Image Anal*, 5(2):111–126. [138](#)
- [Mertens et al., 2013] Mertens, J., De Bruyne, S., Van Damme, N., Smeets, P., Ceelen, W., Troisi, R., Laurent, S., Geboes, K., Peeters, M., Goethals, I., and Van de Wiele, C. (2013). Standardized added metabolic activity (SAM) in 18F-FDG PET assessment of treatment response in colorectal liver metastases. *Eur J Nucl Med Mol Imaging*, 40(8):1214–1222. [24](#)
- [Mertens et al., 2012] Mertens, J., Dobbeleir, A., Ham, H., D’Asseler, Y., Goethals, I., and Van de Wiele, C. (2012). Standardized added metabolic activity (SAM): a partial volume independent marker of total lesion glycolysis in liver metastases. *Eur J Nucl Med Mol Imaging*, 39(9):1441–1448. [24](#)
- [Metzger-Filho et al., 2012] Metzger-Filho, O., Tutt, A., de Azambuja, E., Saini, K. S., Viale, G., Loi, S., Bradbury, I., Bliss, J. M., Azim, H. A., Ellis, P., Di Leo, A., Baselga, J., Sotiriou, C., and Piccart-Gebhart, M. (2012). Dissecting the heterogeneity of triple-negative breast cancer. *J Clin Oncol*, 30(15):1879–1887. [119](#)

- [Miwa et al., 2014] Miwa, K., Inubushi, M., Wagatsuma, K., Nagao, M., Murata, T., Koyama, M., Koizumi, M., and Sasaki, M. (2014). FDG uptake heterogeneity evaluated by fractal analysis improves the differential diagnosis of pulmonary nodules. *Eur J Radiol*, 83(4):715–719. [26](#)
- [Morgan and Bray, 1994] Morgan, D. J. and Bray, K. M. (1994). Lean body mass as a predictor of drug dosage. Implications for drug therapy. *Clin Pharmacokinet*, 26(4):292–307. [18](#)
- [Nestle et al., 2005] Nestle, U., Kremp, S., Schaefer-Schuler, A., Sebastian-Welsch, C., Hellwig, D., Rube, C., and Kirsch, C.-M. (2005). Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. *J Nucl Med*, 46(8):1342–1348. [19](#), [20](#)
- [Nestle et al., 2007] Nestle, U., Schaefer-Schuler, A., Kremp, S., Groeschel, A., Hellwig, D., Rube, C., and Kirsch, C.-M. (2007). Target volume definition for 18F-FDG PET-positive lymph nodes in radiotherapy of patients with non-small cell lung cancer. *Eur J Nucl Med Mol Imaging*, 34(4):453–462. [19](#)
- [Nioche et al., 2013] Nioche, C., Soret, M., Gontier, E., Lahutte, M., Dutertre, G., Dulou, R., Capelle, L., Guillemin, R., Foehrenbach, H., and Buvat, I. (2013). Evaluation of quantitative criteria for glioma grading with static and dynamic 18F-FDopa PET/CT. *Clin Nucl Med*, 38(2):81–87. [130](#)
- [Nowell, 1976] Nowell, P. C. (1976). The clonal evolution of tumor cell populations. *Science*, 194(4260):23–28. [34](#)
- [O'Connor et al., 2015] O'Connor, J.P., R., C.J., W., J.C., C., R.A., P., G.J., J., and A. (2015). Imaging intratumor heterogeneity: role in therapy response, resistance, and clinical outcome. *Clin Cancer Res*, 21(2):249–257. [35](#)
- [Oh et al., 2014] Oh, J. S., Kang, B. C., Roh, J.-L., Kim, J. S., Cho, K.-J., Lee, S.-W., Kim, S.-B., Choi, S.-H., Nam, S. Y., and Kim, S. Y. (2014). Intratumor textural heterogeneity on pretreatment 18F-FDG PET images predicts response and survival after chemoradiotherapy for hypopharyngeal cancer. *Ann Surg Oncol*, 22(8):2746–2754. [46](#), [54](#)
- [Orlhac et al., 2014] Orlhac, F., Soussan, M., Maisonneuve, J.-A., Garcia, C. A., Vanderlinden, B., and Buvat, I. (2014). Tumor texture analysis in 18F-FDG PET: relationships between texture parameters, histogram indices, standardized uptake values, metabolic volumes, and total lesion glycolysis. *J Nucl Med*, 55(3):414–422. [46](#)
- [Patlak and Blasberg, 1985] Patlak, C. S. and Blasberg, R. G. (1985). Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. Generalizations. *J Cereb Blood Flow Metab*, 5(4):584–590. [16](#)
- [Phair, 1997] Phair, R. D. (1997). Development of kinetic models in the nonlinear world of molecular cell biology. *Metabolism*, 46(12):1489–1495. [7](#)
- [Pépin et al., 2014] Pépin, A., Daouk, J., Bailly, P., Hapdey, S., and Meyer, M.-E. (2014). Management of respiratory motion in PET/computed tomography: the state of the art. *Nucl Med Commun*, 35(2):113–122. [19](#)
- [Pyka et al., 2015] Pyka, T., Bundschuh, R. A., Andratschke, N., Mayer, B., Specht, H. M., Papp, L., Zsótér, N., and Essler, M. (2015). Textural features in pre-treatment [F18]-FDG-PET/CT are correlated with risk of local recurrence and disease-specific survival in early stage NSCLC patients receiving primary stereotactic radiation therapy. *Radiat Oncol*, 10(1):100. [23](#), [47](#), [54](#), [126](#), [159](#)
- [Rajkumar et al., 2015] Rajkumar, V., Goh, V., Siddique, M., Robson, M., Boxer, G., Pedley, R. B., and Cook, G. J. R. (2015). Texture analysis of (125)I-A5b7 anti-CEA antibody SPECT differentiates metastatic colorectal cancer model phenotypes and anti-vascular therapy response. *Br J Cancer*, 112(12):1882–1887. [150](#)

- [Richards et al., 2011] Richards, C. H., Mohammed, Z., Qayyum, T., Horgan, P. G., and McMillan, D. C. (2011). The prognostic value of histological tumor necrosis in solid organ malignant disease: a systematic review. *Future Oncol*, 7(10):1223–1235. [126](#)
- [Ruifrok and Johnston, 2001] Ruifrok, A. and Johnston, D. (2001). Quantification of histochemical staining by color deconvolution. *Anal Quant Cytol Histol*, 23(4):291–299. [138](#)
- [Shackleton et al., 2009] Shackleton, M., Quintana, E., Fearon, E. R., and Morrison, S. J. (2009). Heterogeneity in cancer: cancer stem cells versus clonal evolution. *Cell*, 138(5):822–829. [34](#)
- [Shepherd et al., 2012] Shepherd, T., Teras, M., Beichel, R., Boellaard, R., Bruynooghe, M., Dicken, V., Gooding, M., Julyan, P., Lee, J., Lefevre, S., Mix, M., Naranjo, V., Wu, X., Zaidi, H., Zeng, Z., and Minn, H. (2012). Comparative study with new accuracy metrics for target volume contouring in PET image guided radiation therapy. *IEEE Trans Med Imaging*, 31(11):2006–2024. [20](#)
- [Shepp and Vardi, 1982] Shepp, L. A. and Vardi, Y. (1982). Maximum likelihood reconstruction for emission tomography. *IEEE Trans Med Imaging*, 1(2):113–122. [11](#)
- [Smith, 1999] Smith, T. A. (1999). Facilitative glucose transporter expression in human cancer tissue. *Br J Biomed Sci*, 56(4):285–292. [12](#)
- [Soussan et al., 2014] Soussan, M., Orlhac, F., Boubaya, M., Zelek, L., Ziolk, M., Eder, V., and Buvat, I. (2014). Relationship between tumor heterogeneity measured on FDG-PET/CT and pathological prognostic factors in invasive breast cancer. *PloS One*, 9(4):e94017. [46](#)
- [Soussan et al., 2013] Soussan, M., Pop, G., and Eder, V. (2013). Current applications of PET/CT. *Rev Prat*, 63(9):1199–1207. [13](#), [27](#)
- [Sugawara et al., 1999] Sugawara, Y., Zasadny, K. R., Neuhoff, A. W., and Wahl, R. L. (1999). Reevaluation of the standardized uptake value for FDG: variations with body weight and methods for correction. *Radiology*, 213(2):521–525. [18](#)
- [Sundaram et al., 2004] Sundaram, S. K., Freedman, N. M. T., Carrasquillo, J. A., Carson, J. M., Whatley, M., Libutti, S. K., Sellers, D., and Bacharach, S. L. (2004). Simplified kinetic analysis of tumor 18F-FDG uptake: a dynamic approach. *J Nucl Med*, 45(8):1328–1333. [17](#)
- [Szczyński et al., 2009] Szczyński, P. M., Strzelecki, M., Materka, A., and Klepaczko, A. (2009). MaZda—A software package for image texture analysis. *Comp Methods Programs Biomed*, 94(1):66–76. [166](#)
- [Tan et al., 2013a] Tan, S., Kligerman, S., Chen, W., Lu, M., Kim, G., Feigenberg, S., D’Souza, W. D., Suntharalingam, M., and Lu, W. (2013a). Spatial-temporal [18F]FDG-PET features for predicting pathologic response of esophageal cancer to neoadjuvant chemoradiation therapy. *Int J Radiat Oncol Biol Phys*, 85(5):1375–1382. [45](#), [52](#), [55](#), [63](#)
- [Tan et al., 2013b] Tan, S., Zhang, H., Zhang, Y., Chen, W., D’Souza, W. D., and Lu, W. (2013b). Predicting pathologic tumor response to chemoradiotherapy with histogram distances characterizing longitudinal changes in 18F-FDG uptake patterns. *Med Phys*, 40(10):101707. [45](#)
- [Tarantola et al., 2003] Tarantola, G., Zito, F., and Gerundini, P. (2003). PET instrumentation and reconstruction algorithms in whole-body applications. *J Nucl Med*, 44(5):756–769. [8](#)
- [Tateishi et al., 2015] Tateishi, U., Tatsumi, M., Terauchi, T., Ando, K., Niitsu, N., Kim, W. S., Suh, C., Ogura, M., and Tobinai, K. (2015). Prognostic significance of metabolic tumor burden by positron emission tomography/computed tomography in patients with relapsed/refractory diffuse large B-cell lymphoma. *Cancer Sci*, 106(2):186–193. [25](#)
- [Therasse et al., 2000] Therasse, P., Arbuck, S. G., Eisenhauer, E. A., Wanders, J., Kaplan, R. S., Rubinstein, L., Verweij, J., Van Glabbeke, M., van Oosterom, A. T., Christian, M. C., and Gwyther, S. G. (2000). New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst*, 92(3):205–216. [29](#)

- [Thevenaz et al., 1998] Thevenaz, P., Ruttimann, U. E., and Unser, M. (1998). A pyramid approach to subpixel registration based on intensity. *IEEE Trans Image Process*, 7(1):27–41. [138](#)
- [Thibault et al., 2009] Thibault, G., Fertil, B., Navarro, C., Pereira, S., Cau, P., Levy, N., Sequeira, J., and Mari, J. J. (2009). Texture indexes and gray level size zone matrix application to cell nuclei classification. *Pattern Recognition Inf Process*, 140–145. [42](#)
- [Tixier et al., 2014a] Tixier, F., Groves, A. M., Goh, V., Hatt, M., Ingrand, P., Le Rest, C. C., and Visvikis, D. (2014a). Correlation of intra-tumor 18F-FDG uptake heterogeneity indices with perfusion CT derived parameters in colorectal cancer. *PloS One*, 9(6):e99567. [47](#), [48](#), [70](#)
- [Tixier et al., 2012] Tixier, F., Hatt, M., Le Rest, C. C., Le Pogam, A., Corcos, L., and Visvikis, D. (2012). Reproducibility of tumor uptake heterogeneity characterization through textural feature analysis in 18F-FDG PET. *J Nucl Med*, 53(5):693–700. [45](#), [48](#), [50](#), [51](#), [68](#), [92](#), [158](#), [159](#)
- [Tixier et al., 2014b] Tixier, F., Hatt, M., Valla, C., Fleury, V., Lamour, C., Ezzouhri, S., Ingrand, P., Perdrisot, R., Visvikis, D., and Le Rest, C. C. (2014b). Visual versus quantitative assessment of intratumor 18F-FDG PET uptake heterogeneity: prognostic value in non-small cell lung cancer. *J Nucl Med*, 55(8):1235–1241. [23](#), [46](#), [49](#), [54](#), [55](#), [56](#), [91](#), [112](#), [158](#), [159](#), [160](#)
- [Tixier et al., 2011] Tixier, F., Le Rest, C. C., Hatt, M., Albarghach, N., Pradier, O., Metges, J.-P., Corcos, L., and Visvikis, D. (2011). Intratumor heterogeneity characterized by textural features on baseline 18F-FDG PET images predicts response to concomitant radiochemotherapy in esophageal cancer. *J Nucl Med*, 52(3):369–378. [23](#), [45](#), [48](#), [54](#), [55](#), [56](#), [59](#), [61](#), [66](#), [70](#), [91](#), [92](#), [157](#), [159](#), [160](#)
- [Uematsu, 2011] Uematsu, T. (2011). MR imaging of triple-negative breast cancer. *Breast Cancer*, 18(3):161–164. [119](#)
- [Vaidya et al., 2012] Vaidya, M., Creach, K. M., Frye, J., Dehdashti, F., Bradley, J. D., and El Naqa, I. (2012). Combined PET/CT image characteristics for radiotherapy tumor response in lung cancer. *Radiother Oncol*, 102(2):239–245. [25](#), [45](#), [63](#)
- [Valli  res et al., 2015] Valli  res, M., Freeman, C. R., Skamene, S. R., and El Naqa, I. (2015). A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. *Phys Med Biol*, 60(14):5471–5496. [165](#)
- [van G  mez L  pez et al., 2014] van G  mez L  pez, O., Garc  a Vicente, A. M., Honguero Mart  nez, A. F., Soriano Castrej  n, A. M., Jim  nez Londo  o, G. A., Udias, J. M., and Le  n Atance, P. (2014). Heterogeneity in [18F]fluorodeoxyglucose positron emission tomography/computed tomography of non-small cell lung carcinoma and its relationship to metabolic parameters and pathologic staging. *Mol Imaging*, 13:1–12. [47](#), [125](#), [159](#)
- [van Velden et al., 2011] van Velden, F. H. P., Cheebsumon, P., Yaqub, M., Smit, E. F., Hoekstra, O. S., Lammertsma, A. A., and Boellaard, R. (2011). Evaluation of a cumulative SUV-volume histogram method for parameterizing heterogeneous intratumoural FDG uptake in non-small cell lung cancer PET studies. *Eur J Nucl Med Mol Imaging*, 38(9):1636–1647. [25](#)
- [van Velden et al., 2014] van Velden, F. H. P., Nissen, I. A., Hayes, W., Velasquez, L. M., Hoekstra, O. S., and Boellaard, R. (2014). Effects of reusing baseline volumes of interest by applying (non-)rigid image registration on positron emission tomography response assessments. *PloS One*, 9(1):e87167. [25](#)
- [Vanderhoek et al., 2012] Vanderhoek, M., Perlman, S. B., and Jeraj, R. (2012). Impact of the definition of peak standardized uptake value on quantification of treatment response. *J Nucl Med*, 53(1):4–11. [21](#)
- [Vauclin et al., 2010] Vauclin, S., Gardin, I., Doyeux, K., Hapdey, S., Edet-Sanson, A., and Vera, P. (2010). Segmentation des images TEP au 18F-FDG. Principe et revue de la litt  rature. *Med Nucl*, 34(6):358–369. [20](#)

- [Visvikis et al., 2012] Visvikis, D., Hatt, M., Tixier, F., and Cheze Le Rest, C. (2012). The age of reason for FDG PET image-derived indices. *Eur J Nucl Med Mol Imaging*, 39(11):1670–1672. [44](#)
- [Vogelstein et al., 2013] Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A., and Kinzler, K. W. (2013). Cancer genome landscapes. *Science*, 339(6127):1546–1558. [35](#)
- [Vuong et al., 2014] Vuong, D., Simpson, P. T., Green, B., Cummings, M. C., and Lakhani, S. R. (2014). Molecular classification of breast cancer. *Virchows Arch*, 465(1):1–14. [115](#)
- [Wahl et al., 2009] Wahl, R. L., Jacene, H., Kasamon, Y., and Lodge, M. A. (2009). From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. *J Nucl Med*, 50 Suppl 1:122S–50S. [21](#), [30](#)
- [Warburg, 1956] Warburg, O. (1956). On respiratory impairment in cancer cells. *Science*, 124(3215):269–270. [12](#)
- [Watabe et al., 2011] Watabe, T., Tatsumi, M., Watabe, H., Isohashi, K., Kato, H., Yanagawa, M., Shimosegawa, E., and Hatazawa, J. (2011). Intratumoral heterogeneity of F-18 FDG uptake differentiates between gastrointestinal stromal tumors and abdominal malignant lymphomas on PET/CT. *Ann Nucl Med*, 26(3):222–227. [25](#)
- [Waterton and Pykkanen, 2012] Waterton, J. C. and Pykkanen, L. (2012). Qualification of imaging biomarkers for oncology drug development. *Eur J Cancer*, 48(4):409–415. [166](#)
- [Watson, 2000] Watson, C. (2000). New, faster, image-based scatter correction for 3D PET. *IEEE Trans Nucl Sci*, 47(4):1587–1594. [137](#)
- [WHO, 1979] WHO (1979). *WHO handbook for reporting results of cancer treatment*, volume 48. World Health Organization, Genève. [29](#)
- [Willaime et al., 2013] Willaime, J. M. Y., Turkheimer, F. E., Kenny, L. M., and Aboagye, E. O. (2013). Quantification of intra-tumour cell proliferation heterogeneity using imaging descriptors of 18F-fluorothymidine-positron emission tomography. *Phys Med Biol*, 58(2):187–203. [46](#), [48](#), [54](#), [62](#), [92](#), [101](#)
- [Wong et al., 2002] Wong, K.-P., Feng, D., Meikle, S., and Fulham, M. (2002). Segmentation of dynamic PET images using cluster analysis. *IEEE Trans Nucl Sci*, 49(1):200–207. [21](#)
- [Woodard et al., 1975] Woodard, H. Q., Bigler, R. E., Freed, B., and Russ, G. (1975). Expression of tissue isotope distribution. *J Nucl Med*, 16(10):958–959. [17](#)
- [Xu et al., 2004] Xu, D.-H., Kurani, A. S., Furst, J. D., and Raicu, D. S. (2004). Run-length encoding for volumetric texture. *The 4th IASTED International Conference on Visualization, Imaging and Image Processing: VIP; September 6–8, 2004; Marbella, Espagne*. [41](#)
- [Xu et al., 2014] Xu, R., Kido, S., Suga, K., Hirano, Y., Tachibana, R., Muramatsu, K., Chagawa, K., and Tanaka, S. (2014). Texture analysis on (18)F-FDG PET/CT images to differentiate malignant and benign bone and soft-tissue lesions. *Ann Nucl Med*, 28(9):926–935. [47](#), [54](#), [92](#), [101](#), [159](#)
- [Yan et al., 2015] Yan, J., Lim, J. C.-S., Loi, H. Y., Khor, L. K., Sinha, A. K., Quek, S. T., Tham, I. W., and Townsend, D. W. (2015). Impact of image reconstruction settings on texture features in 18F-FDG PET. *J Nucl Med (sous presse)*. [158](#), [159](#), [160](#), [162](#)
- [Yang et al., 2013] Yang, F., Thomas, M. A., Dehdashti, F., and Grigsby, P. W. (2013). Temporal analysis of intratumoral metabolic heterogeneity characterized by textural features in cervical cancer. *Eur J Nucl Med Mol Imaging*, 40(5):716–727. [46](#), [70](#), [158](#), [160](#)
- [Yip et al., 2014] Yip, S., McCall, K., Aristophanous, M., Chen, A. B., Aerts, H. J. W. L., and Berbeco, R. (2014). Comparison of texture features derived from static and respiratory-gated PET images in non-small cell lung cancer. *PloS One*, 9(12):e115510. [47](#), [50](#)

- [Young et al., 1999] Young, H., Baum, R., Cremerius, U., Herholz, K., Hoekstra, O., Lammertsma, A. A., Pruim, J., and Price, P. (1999). Measurement of clinical and subclinical tumour response using [18F]-fluorodeoxyglucose and positron emission tomography: review and 1999 EORTC recommendations. European Organization for Research and Treatment of Cancer (EORTC) PET Study Group. *Eur J Cancer*, 35(13):1773–1782. [29](#)
- [Yu et al., 2009a] Yu, H., Caldwell, C., Mah, K., and Mozeg, D. (2009a). Coregistered FDG PET/CT-based textural characterization of head and neck cancer for radiation treatment planning. *IEEE Trans Med Imaging*, 28(3):374–383. [55](#)
- [Yu et al., 2009b] Yu, H., Caldwell, C., Mah, K., Poon, I., Balogh, J., MacKenzie, R., Khaouam, N., and Tirona, R. (2009b). Automated radiation targeting in head-and-neck cancer using region-based texture analysis of PET and CT images. *Int J Rad Oncol Biol Phys*, 75(2):618–625. [45](#), [54](#)
- [Zaidi and Koral, 2004] Zaidi, H. and Koral, K. F. (2004). Scatter modelling and compensation in emission tomography. *Eur J Nucl Med Mol Imaging*, 31(5):761–782. [19](#)
- [Zeng, 2001] Zeng, G. L. (2001). Image reconstruction — a tutorial. *Comput Med Imaging Graph*, 25(2):97–103. [11](#)
- [Zhang et al., 2014] Zhang, H., Tan, S., Chen, W., Kligerman, S., Kim, G., D’Souza, W. D., Suntharalingam, M., and Lu, W. (2014). Modeling pathologic response of esophageal cancer to chemoradiation therapy using spatial-temporal 18F-FDG PET features, clinical parameters, and demographics. *Int J Radiat Oncol Biol Phys*, 88(1):195–203. [46](#)
- [Zhu and Shim, 2011] Zhu, A. and Shim, H. (2011). Current molecular imaging positron emitting radiotracers in oncology. *Nucl Med Mol Imaging*, 45(1):1–14. [14](#)
- [Zincirkeser et al., 2007] Zincirkeser, S., Şahin, E., Halac, M., and Sager, S. (2007). Standardized uptake values of normal organs on 18F-Fluorodeoxyglucose positron emission tomography and computed tomography imaging. *J Int Med Res*, 35(2):231–236. [164](#)
- [Zips et al., 2012] Zips, D., Zöphel, K., Abolmaali, N., Perrin, R., Abramjuk, A., Haase, R., Appold, S., Steinbach, J., Kotzerke, J., and Baumann, M. (2012). Exploratory prospective trial of hypoxia-specific PET imaging during radiochemotherapy in patients with locally advanced head-and-neck cancer. *Radiother Oncol*, 105(1):21–28. [28](#)

Annexe A

Liste des abréviations

18F-FDG : 2-desoxy-2-18F-fluoro-D-glucose

AMM : Autorisation de Mise sur le Marché

AR : image AutoRadiographique

ARR : image AutoRadiographique Ré-échantillonnée

ARRL : image AutoRadiographique Ré-échantillonnée et Lissée

AUC : Aire sous la Courbe ROC

BSA : *Body Surface Area*

BTV : *Biologically Tumor sub-Volume*

CCM : Cancer Colorectal Métastatique

CM : Cancer Mammaire

CO : Matrice de Co-Occurrence

COV : COefficient de Variation

CPNPC : Cancer Pulmonaire Non à Petites Cellules

CSH : *Cumulative SUV-volume Histogram*

CTV : *Clinical Target Volume*

D : Taux de discrétisation

DA : Discrétisation Absolue (équation [3.5](#))

DR : Discrétisation Relative (équation [3.2](#))

DR+1 : Discrétisation Relative avec un +1 au dénominateur (équation [3.4](#))

EC : Écart-type

EORTC : *European Organization for Research and Treatment of Cancer*

ESKA : *Extension of the Simplified Kinetic Analysis*

EVP : Effet de Volume Partiel

FLAB : *Fuzzy Locally Adaptive Bayesian*

GLNUr : *Gray-Level Non-Uniformity for run*

GLNUz : *Gray-Level Non-Uniformity for zone*

GLRLM : *Gray-Level Run Length Matrix* ou matrice des longueurs des séries homogènes

GLZLM : *Gray-Level Zone Length Matrix* ou matrice des longueurs des zones homogènes

GTV : *Gross Tumor Volume*

H : image après marquage à l'Hématoxyline
 H&E : marquage des coupes histologiques à l'Hématoxyline et Éosine
 HGRE : *High Gray-level Run Emphasis*
 HGZE : *High Gray-level Zone Emphasis*
 HL : image après marquage à l'Hématoxyline et Lissage

ICC : *Intra-class Correlation Coefficient*
 IRM : Imagerie par Résonance Magnétique
 IVH : *Intensity Volume Histogram*

LBM : *Lean Body Mass*
 LGRE : *Low Gray-level Run Emphasis*
 LGZE : *Low Gray-level Zone Emphasis*
 LRE : *Long-Run Emphasis*
 LRHGE : *Long-Run High Gray-level Emphasis*
 LRLGE : *Long-Run Low Gray-level Emphasis*
 LZE : *Long-Zone Emphasis*
 LZHGE : *Long-Zone High Gray-level Emphasis*
 LZLGE : *Long-Zone Low Gray-level Emphasis*

MLN : *Maximum Likelihood Expectation Maximization*

NGLDM : *Neighborhood Gray-Level Different Matrix* ou matrice de différences des niveaux de gris

OMS : Organisation Mondiale de la Santé
 OSEM : *Ordered Subset Expectation Maximization*

PERCIST : *Positron Emission tomography Response Criteria in Solid Tumors*
 PTV : *Planning Target Volume*

r : Coefficient de corrélation de Pearson
 RAMLA : *Row Action Maximum Likelihood Algorithm*
 RECIST : *Response Evaluation Criteria in Solid Tumors*
 RLNU : *Run Length Non-Uniformity*
 ROC : *Receiver Operating Characteristic*
 RP : *Run Percentage*
 r_s : Coefficient de corrélation de Spearman

SAM : *Standardized Added Metabolic activity*
 SKA-M : *Simplified Kinetic Analysis using Multiple time points*
 SKM : *Simplified Kinetic Method*
 SRE : *Short-Run Emphasis*
 SRHGE : *Short-Run High Gray-level Emphasis*
 SRLGE : *Short-Run Low Gray-level Emphasis*
 SUV : *Standardized Uptake Value* ou valeur de fixation normalisée
 SZE : *Short-Zone Emphasis*
 SZHGE : *Short-Zone High Gray-level Emphasis*
 SZLGE : *Short-Zone Low Gray-level Emphasis*

TDM : Tomodensitométrie
TEP : Tomographie par Émission de Positons
TLG : *Total Lesion Glycolysis* ou glycolyse de la lésion
TN : tumeur mammaire Triple-Négative
TOF : *Time Of Flight* ou Temps de Vol

VM : Volume Métabolique
VOI : VOlume d'Intérêt
VOI-C : VOlume d'Intérêt reporté dans le tissu Controlatéral
VOI-F : VOlume d'Intérêt reporté dans le Foie
VOI-S : VOlume d'Intérêt Sphérique dans le fantôme
VOI-T : VOlume d'Intérêt Tumoral

ZLNU : *Zone Length Non-Uniformity*
ZP : *Zone Percentage*

Annexe B

Liste des publications

Articles dans des revues avec comité de lecture (acceptés)

- *Tumor texture analysis in PET: where do we stand?* I. Buvat, **F. Orlhac**, M. Soussan. J. Nucl. Med. 56: sous presse, 2015.
- *Tumor texture analysis in 18F-FDG PET: relationships between texture parameters, histogram indices, Standardized Uptake Values, metabolic volumes and total lesion glycolysis.* **F. Orlhac**, M. Soussan, JA. Maisonobe, CA. Garcia, B. Vanderlinden, I. Buvat. J. Nucl. Med. 55: 414–422, 2014.
- *Relationship between tumor heterogeneity measured on FDG-PET/CT and pathological prognostic factors in invasive breast cancer.* M. Soussan, **F. Orlhac**, M. Boubaya, L. Zelek, M. Ziol, V. Eder, I. Buvat. PloS ONE. 9:e94017, 2014.
- *Volume-based indices on FDG-PET/CT in locally advanced non-small-cell lung cancer: prediction of histological response after induction chemotherapy.* M. Soussan, J. Cyrta, C. Pouliquen, K. Chouahnia, **F. Orlhac**, E. Martinod, V. Eder, JF. Morère, I. Buvat. Radiology. 272:875–884, 2014.

Communications orales avec actes (acceptées)

- *Enhanced textural indices distinguish tissue types in FDG PET.* **F. Orlhac**, M. Soussan, K. Chouahnia, E. Martinod, I. Buvat. Society of Nuclear Medicine and Molecular Imaging annual meeting, 6–10 juin 2015, Baltimore, USA. J. Nucl. Med. 56:430, 2015.
- *Co-analysis of textural pattern and SUVpeak in F18-FDOPA PET/CT significantly improves glioma staging.* C. Nioche, **F. Orlhac**, M. Soret, E. Gontier, I. Buvat. Society of Nuclear Medicine and Molecular Imaging annual meeting, 6–10 juin 2015, Baltimore, USA. J. Nucl. Med. 56:327, 2015.
- *Can texture indices derived from PET images differentiate tissue types?* **F. Orlhac**, M. Soussan, I. Buvat. Colloque Recherche en Imagerie et Technologies pour la Santé, 25–27 mars 2015, Dourdan. Actes des journées RITS:96–97.
- *Contribution of tissue textural pattern and conventional index to glioma staging in FDopa-PET/CT.* C. Nioche, **F. Orlhac**, M. Soret, E. Gontier, I. Buvat. Colloque Recherche en Imagerie et Technologies pour la Santé, 25–27 mars 2015, Dourdan. Actes des journées RITS:50–51.

- *Do texture indices derived from PET improve tumor characterization in non-small cell lung cancer (NSCLC) patients ?* **F. Orlhac**, M. Soussan, V. Eder, I. Buvat. Society of Nuclear Medicine and Molecular Imaging annual meeting, 7–11 juin 2014, Saint Louis, USA. J. Nucl. Med. 55:453, 2014.
- *Texture indices for tumor characterization in FDG PET : robustness and relationship with SUV and metabolic volumes.* **F. Orlhac**, M. Soussan, JA. Maisonobe, CA. Garcia, B. Vanderlinden, I. Buvat. European Association of Nuclear Medicine annual congress, 19–23 octobre 2013, Lyon, France. Eur. J. Nucl. Med. Mol. Imaging. 40:OP 518, 2013.
- *FDG-PET/CT textural features in triple negative breast tumors.* M. Soussan, **F. Orlhac**, M. Bou-baya, L. Zelek, M. Zioli, V. Eder, I. Buvat. European Association of Nuclear Medicine annual congress, 19–23 octobre 2013, Lyon, France. Eur. J. Nucl. Med. Mol. Imaging. 40:OP 543, 2013.

Communications affichées avec actes (acceptées)

- *Entropy characterizes intratumoral hypometabolism in FDG PET.* **F. Orlhac**, M. Soussan, S. Djelbani, J. Tordjmann, I. Buvat. Society of Nuclear Medicine and Molecular Imaging annual meeting, 6–10 juin 2015, Baltimore, USA. J. Nucl. Med. 56:1779, 2015.
- *Entropy of tumor FDG uptake measured in-vivo from PET images reflects the tumor entropy seen in autoradiography.* **F. Orlhac**, B. Thézé, R. Boisgard, I. Buvat. Society of Nuclear Medicine and Molecular Imaging annual meeting, 6–10 juin 2015, Baltimore, USA. J. Nucl. Med. 56:1777, 2015.
- *Apport de la texture tissulaire et des indices conventionnels en TEP/TDM à la 18FDopa pour l'évaluation du grade tumoral des gliomes.* C. Nioche, **F. Orlhac**, M. Soret, E. Gontier, I. Buvat. Premières journées francophones de médecine nucléaire, 28–31 mai 2015, La Rochelle. Médecine Nucléaire 39(3):226–227, 2015.
- *Comparison of tumor textural features in PET and CT images in non-small cell lung cancer (NSCLC) patients.* **F. Orlhac**, M. Soussan, V. Eder, I. Buvat. Society of Nuclear Medicine and Molecular Imaging annual meeting, 7–11 juin 2014, Saint Louis, USA. J. Nucl. Med. 55:1610, 2014.
- *Glioma grading in 18F-Dopa PET/CT: comparison between textural feature, shape index and SUV.* **F. Orlhac**, C. Nioche, M. Soret, E. Gontier, I. Buvat. Society of Nuclear Medicine and Molecular Imaging annual meeting, 7–11 juin 2014, Saint Louis, USA. J. Nucl. Med. 55:1895, 2014.
- *Correlation between FDG-PET/CT textural features and molecular subtypes of breast cancer.* M. Soussan, **F. Orlhac**, L. Zelek, M. Zioli, V. Eder, I. Buvat. Society of Nuclear Medicine and Molecular Imaging annual meeting, 8–12 juin 2013, Vancouver, Canada. J. Nucl. Med. 54:1446, 2013.

Conférence invitée

- *Méthodologie de l'évaluation thérapeutique : un monde parfait. Vers une généralisation des mesures semi-quantitatives : caractérisation de la texture tumorale.* **F. Orlhac**, I. Buvat. Enseignement Post-Universitaire « TEP : aspects physiques et quantitatifs », 20 juin 2014, Paris.

Articles en cours de soumission

- *^{18}F -FDG PET-derived textural indices reflect tissue-specific information in non-small cell lung cancer.* **F. Orlhac**, M. Soussan, K. Chouahnia, E. Martinod, I. Buvat.
- *Interpretation of texture indices in PET.* **F. Orlhac**, M. Soussan, G. Pop, I. Buvat.
- *Multi-scale texture analysis: from ^{18}F -FDG PET images to histological sections.* **F. Orlhac**, B. Thézé, M. Soussan, R. Boisgard, I. Buvat.